

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 09:28:05 : Search time 7690.28 Seconds
(without alignments)
10963.461 Million cell updates/sec

Title: US-10-712-479-3
Perfect score: 1740
Sequence: 1 atggctagcagtcacatgc.....gccatcacccatcacatcac 1740

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.hcg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1595.4	91.1	9547	AB049091	AB049091 Hepatitis
2	1583.8	91.0	9548	D89815	D89815 Hepatitis C
3	1582.2	90.9	9374	AF207753	AF207753 Hepatitis
4	1574.2	90.5	9379	AF165059	AF165059 Hepatitis
5	1572.6	90.4	9379	AF165060	AF165060 Hepatitis
6	1572.4	90.4	1773	AB057599	AB057599 Hepatitis
7	1566.6	90.0	9375	AF207754	AF207754 Hepatitis
8	1563	89.8	9616	AB049088	AB049088 Hepatitis
9	1561.4	89.7	9377	AF207763	AF207763 Hepatitis
10	1559.8	89.6	9379	AF207767	AF207767 Hepatitis
11	1558.2	89.6	9456	AF165059	AF165059 Hepatitis
12	1556.6	89.5	9448	HPCRNA	HPCRNA
13	1554	89.3	9379	D10934	D10934 Hepatitis C
14	1553.8	89.3	1773	AF165062	AF165062 Hepatitis C
15	1553.4	89.3	1773	AB057601	AB057601 Hepatitis
16	1553.4	89.3	1743	AX686788	AX686788 Sequence
17	1552.2	89.3	1743	BD010839	BD010839 HCV poly
18	1551.8	89.2	9580	AF054250	AF054250 Hepatitis C
19	1551.8	89.2	9414	M96362	M96362 Hepatitis C
20	1551.8	89.2	9415	HCU16362	HCU16362 Hepatitis C

20	1551.8	89.2	9460	14	AB049097	AB049097 Hepatitis
21	1551.8	89.2	9472	6	AR301390	AR301390 Sequence
22	1549.8	89.1	9435	14	AB049093	AB049093 Hepatitis
23	1549.2	89.0	9377	14	AF207756	AF207756 Hepatitis
24	1549.2	89.0	9535	14	D85516	D85516 Hepatitis C
25	1548.6	89.0	5137	14	AY044867	AY044867 Hepatitis
26	1548.6	89.0	9587	14	AB080299	AB080299 Hepatitis
27	1548.4	89.0	1773	14	AB031322	AB031322 Hepatitis
28	1548.4	89.0	1773	14	AB057680	AB057680 Hepatitis
29	1547	88.9	9460	14	HPC3491	D10750 Hepatitis C
30	1547	88.9	9595	6	AR119832	AR119832 Sequence
31	1547	88.9	9595	6	CO819761	CO819761 Sequence
32	1547	88.9	9595	14	AF054247	AF054247 Hepatitis
33	1546.8	88.9	9573	14	AB049090	AB049090 Hepatitis
34	1545.4	88.8	9595	14	AF054248	AF054248 Hepatitis
35	1545.2	88.8	1773	14	AB057688	AB057688 Hepatitis
36	1543.8	88.7	9379	14	AF207758	AF207758 Hepatitis
37	1543.8	88.7	9379	14	AF207772	AF207772 Hepatitis
38	1543.6	88.7	9587	14	AF333324	AF333324 Hepatitis
39	1543.6	88.7	1773	14	AB057602	AB057602 Hepatitis
40	1542.8	88.7	9379	14	AF165061	AF165061 Hepatitis
41	1542.2	88.6	9379	14	AF165058	AF165058 Hepatitis
42	1542.2	88.6	9446	14	HCU01214	U01214 Hepatitis C
43	1542.2	88.6	9596	14	AF054249	AF054249 Hepatitis
44	1541.8	88.6	1779	6	AX593183	AX593183 Sequence
45	1540.8	88.6	9359	14	AF313916	AF313916 Hepatitis

ALIGNMENTS

RESULT 1	AB049091	AB049091	Hepatitis C virus gene for polyprotein, complete cds,	9547 bp	RNA	linear	VRL 22-AUG-2002
LOCUS	AB049091	AB049091	isolate:HCV142.				
DEFINITION	AB049091	AB049091.1	GI:11559448				
ACCESSION	AB049091	AB049091.1	GI:11559448				
VERSION	AB049091	AB049091.1	GI:11559448				
KEYWORDS	AB049091	AB049091.1	GI:11559448				
SOURCE	AB049091	AB049091.1	GI:11559448				
ORGANISM	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB049091.1	GI:11559448				
AUTHORS	AB049091	AB049091.1	GI:11559448				
TITLE	AB049091	AB049091.1	GI:11559448				
JOURNAL	AB049091	AB049091.1	GI:11559448				
PUBMED	AB049091	AB049091.1	GI:11559448				
REFERENCE	AB049091	AB04					

/db_xref="GI:11559449"
/translation="MSTNRPKPKRXTKRNTRRPODVKFGGGQIVGGVYLIPRGRPL
GVRAATRSRSPGRQPIPKARQPEGRAMAQPGVPMPLYGNEGILWAGNLSLRG
SRPSWGDPDRRRNRNLKVDLTJTCGFADLMGYIPLVGAPLGAALAHAGVRVLED
GVNTAGNLPCCSFSIPELLALSCLTTPASAYEVNVSGIYHVNTDCSNSIYVEAAD
VIMHAGVPCVRNNSSRWALTPILAAHNSVPTTIRRHVDLLVGTAAFCSAWY
VGLDCGSFLLTSQLPTFSRPRHETVQSCNSLYPHVTHGRMAWDMNWSPTAALV
SOLLRIPOAVMDVAGAHGVLGLAYISVNGWNAKVL1VMLLFAGVGGTHVTGAV
GRSFLRSTPSTGASOKIOLINTNGSHNRKALANCDSLHGTGFLAALFYTHRFNAS
CSPRLASCRSIDTFDQMGPIITYAOPSSSDQRPYCHYAPRCGIIYALFYOYGVPC
PTSPVVGTITDFRGVPTYSWENETDULLNNTRPPQGNWFGCTWNMS"GFKTCCG
TICNVIGAGNNLTCTDCFXHPEATYKCGSGPMLTPRCNVDPYIRLWHYFCTVNF
TIFKVMGVGVEHRLNAACNWRGERCNREDRSELSPLLSTTETWQILPCLSTTL
PALSTGLHLHONVVDVQYLYGVGSAVSVLWKWEYVLLFLLLADARVCACLMMLL
IAQREALENLAVLNAASVAGHGLSPLVFPCCAAWYIKGKLVPCAAYAFYGVMPILL
LILALPRAYANDREMAASCQGVFVGLVLLTSLPYVKVFLAKLIWLOYFIITRAEAB
LQVVPPLNVGRGRDAIILITCAAHPBLIFDITKLLALFGLMLVLAQGITRYPYVR
AQGLIRACMLVRVAGHYVQMAFMKLAALGTYYVYDHLPLRDMWHTGLRDLAVAE
PVFSDEMTKIIITWADTAACDIIISGLPVSAARRGRELIGPADSLGCGWRELLAPIT
AYSOOTRGLGCIITSLTRDKNOVEGVQVSTATOSFLATCVNGVCWTYHAGSK
TLAGPKGITOMYNVDODLVGNQAPPGARSITPCTCGSSDLXLYTRHADVIPVRRG
DSRGLSPRPVSYLKGSSGGLPCPSGHVUGIIFRAAVCTRGVAKAVDFVPEAMET
MRSPVTDNSPPAVPQTFQVAHLHAPTGSKSTKVPAAVAAQTKVLVNLPSVAATL
CFGAVMSKHGTDPMIRTGIRITITGAPITYTYGKFLADGCGSGAYDIIICDECHS
TDSITLIGITVLDQAEAGARLVVLTATAPGSGVTPHPNIEBVALNTGEIIPYOK
ALPIEAIKGBRHLIFCHSKKCDLEAAKLSSIGINAVAYRGLDVSIVPTSGNVVVA
TDALMGTYGDPDSVIDGNTCVTQTVDFSLDPTFTIETTVQDADVSRTORRGRTGR
RGIYFVTPGEPSPGMSDVLCEYDAGCAWELTPAETSVRLRAYLNTPLGIPVQC
DHLEWESVFTGLTHIDAFILSOTQAGDNFPYLVAYOATVCARAQAPPPSWDQMKC
LIRLPLTHGPTPLLYRLGAVQNEVTLTHPVTKYIMACMSADLEVTSTVWLVGGVLA
ALAAVCLTGSWIVGRIILISGRPAVIPDRVLRPEFDEMECEASHLPYIEQGMQLE
QFKQALGLLOTATQAEAAAPVVEKWRLEAFWAKHMNFISGIOVLAGLSTLPGV
PAIASLMAPTASVTSPLTTOSTLLFNILGHWAAQLAPPSAASAFVAGIAGAAVGSII
GLGKVLVDILAGVAGVAGALVAFKVMGSEMPSTEDLVNLLPAILSPGALVGVVCAA
LURRHVGPGEAGVOMNRLIAPSRGNHVSPTHYVPESDAAARVTQILSSLTITLQILK
RLHOMINEDSTPCSGSLWRDWDWICTVLIDFTKWLQSKLLPLRGVYPPFSCQRYK
GWRGDMHTTCCGAQIAGHVXGSMRIIVGPKTCSNTWGTPTPINAYTTGPCTPSP
APNYBRAUMVAAEYEVTRVGDHYITGMTDNIKCPQVPAPEFTEVDVGLRLHR
YAPACKPLREBAVFOVGNLOYLVGSQPEPEPDVAVLTSMLTDPSTIETAKRRL
ARGSPSPMASSASQSLPALKAACTTHHSPLDLIENAILRMQEMGNITRVESN
KVILDSFDPLRAEBDEREVSVPAILLAKSRNFPAMPPIWARPDYNNPLLESWKDPOY
VPPVHGCPPLPTKAPPVPPRRKRTVLTSTVSSALAEATKTFGSSSESAAADST
ATAPDQASDDGKSDVESYSMPLEGPDPULSDGSWSVSEEAETDQVCCSMS
YTWGALITPCAAEESKPIINALSNLLRHNNVATTSRASLRQKVTFDRLVILD
DIYRDVLKEMKASTVKAALLSVEEACKLTPPSAKSKFGYKADKXNLSKATNIH
RSVWKDLLEDTEPTIDITIMAKNEVFCVQPEGRKPARLIIVPDLGVYRVCHOMALYD
VYSTLPQAVMGSSYQFQYSPGORVFLVTKWSKKCPMGFSYDTRCFDSTVTENDIRV
ESIIYOCDLAPEARQAIIRSLTERLYIGGPMTNSKQNGCYPRCHASGVLTSCGNTL
TCKLKATAACRAKLUQDCTMLVNGDDLVIICRSAGTQEDAAANLVRVTEAMTYSPAPG
DPVPEYDLELITSCSSNVSAHDASGRVYILTRDPTPLSLRAWETARHTPVNSWL
GNIIYAPTILWARMILMTHFFSILLAQEQLEKALDQIYVAGCYSIEPLDLPQILERLH
GLSASFSLHSVSPGEINRVASCLRLKGVPLRVWRHRAVSRAKLISQOGRATCKYL
RWAIVTKLKTPIPAASQLDLSGMFVAGYSGDIYHLSLSRARPRWFMLCLILLSVG
GIYLLPNR"
9293. .9331
/note="upstream of poly-U stretch"
9332. .9425
/note="poly(U) length is indistinct"
poly(U) stretch
9426. .9547
/note="3' X tail"

ORIGIN

Query Match 91.1%; Score 1585.4; DB 14; Length 9547;
Best local similarity 95.3%; Pred. No. 0;
Matches 1634; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
Qy 8 GCATGTCATCTCTATATGAGACAGCGCCCTTGATCACACCGTGGCTGCGAGGAAA 67
Db 7512 GCTGCTCAATGCTCTACATGAGACAGCGCCCTTGATCACCGCATGCGTGGCGAGGAAA 7571
Qy 68 GCAAGCTGCCCATCAACCGCGCTGAGCAACTCTTCTGCTGGCTACCATAACTGCTTATT 127

Db 7572 GCAAGCTGCCCATCAACCGCGCTGAGCAACTCTTCTGCTGGCTCACCAACAATGCGGTGTATG 7631
Qy 128 CCACAACATCCCGCAGTCAAGCCTCGGCGACAAGAAGTCACTCTTGTGACAGACTGCAAG 187
Db 7632 CCACAACATCCCGCAGCGCAAGCCTCGGCGACAAGAAGTCACTCTTGTGACAGACTGCAAG 7691
Qy 188 TCTGCAAGCATCATTACCCGGACGTGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTGA 247
Db 7692 TCTGCAAGCATCATTACCCGGACGTGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTGA 7751
Qy 248 AGGCTAAATCTGCTATCTGTAGAAGAAAGCATGCAAGCTGACGCCCGCGCATTTGGCCAAAT 307
Db 7752 AGGCTAAATCTTCTATCTGTAGAAGAAAGCCTGCAAGCTGACGCCCGCGCATTTGGCCAAAT 7811
Qy 308 CCAAAATTTGGCTATGGGCGAAAGGAGCTGCGGAGCCTTCCAGCAGGCGCGTTAAACACA 367
Db 7812 CCAAAATTTGGCTATGGGCGAAAGGAGCCTCGGGAACCTATCCAGCAAGGGCCATTTAAACACA 7871
Qy 368 TCCGCTCCGCTGCGAAGGACTTGTCTGAGGACACTGCACACACCAATTTACAGACACCATCA 427
Db 7872 TCCGCTCCGCTGCGAAGGACTTGTCTGAGGACACTGCACACACCAATTTACAGACACCATCA 7931
Qy 428 TGGCAAAAATAGAGTTTCTGCGTCCAAACAGAGAAAGAGGCGCGCAACACAGCTCGCC 487
Db 7932 TGGCAAAAATAGAGTTTCTGCGTCCAGCCAGAAAGAGGCGCGCAACACAGCTCGCC 7991
Qy 488 TCATCGTATTTCCAGACCTCGGAGTTTGTGTATGCGAAGAGATGGCCCTTTACGACGTGG 547
Db 7992 TTATCGTATTTCCAGACCTCGGAGTTTGTGTATGCGAAGAGATGGCCCTTTATGACGTAG 8051
Qy 548 TTTCCACTCTTCTCAGGCCGTGATGGCTCCCTCATACGGAATTTCCAATCTCTCCTAAGC 607
Db 8052 TTTCCACTCTTCTCAGGCCGTGATGGCTCCCTCATACGGAATTTCCAATCTCTCCTGAC 8111
Qy 608 AGCGGTCGAGTTCTCTGTAATACCTTGGAAAGCAAGAAATGCCCTATGGGCTTTCTCAT 667
Db 8112 AGCGGTCGAGTTCTCTGTAAGACCTTGGAAATGGAAGATGGCCCTATGGGCTTTCTCAT 8171
Qy 668 ATGACACCCGCTGTTTGTGACTCAACCGTCACTAGAAATGACATCCGTTGTGAGGAGTCAA 727
Db 8172 ATGACACCCGTTGTTTGTGACTCAACCGTCACTAGAAATGACATCCGTTGTGAGGAGTCAA 8231
Qy 728 TTTTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCCATTAAGGTTCGTCTACAGAGC 787
Db 8232 TTTTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCCATTAAGGTTCGTCTACAGAGC 8291
Qy 788 GGTCTATGTCGGGGTCCCATGACTTCCAAGGCGCAACTGGCGGCTATCGCCGCT 847
Db 8292 GGTCTTATATCGGGGGCCCCATGACTTCCAAGGCGCAACTGGCGGCTATCGCCGCT 8351
Qy 848 GCGGCGGAGCGGGTGTGCTGACGACTAGCTCGGGTAATACCTCACAATGCTACTTTGAAG 907
Db 8352 GCGGCGGAGCGGGTGTGCTGACGACTAGCTCGGGTAATACCTCACAATGCTACTTTGAAG 841
Qy 908 CCGCTGCAAGCCTGTCGAGCTGCCAAGCTTCCAGGCTGCAAGATGCTGTAATGAGAGAGC 967
Db 8412 CCACTGCAAGCCTGTCGAGCTGCAAGCTTCCAGGCTGCAAGATGCTGTAAGCGGAGAGC 847
Qy 968 ACCTTGTGCTTATCTGTGAAGCGGGGAAACCCAGAGAGCGGCGCAAGCTTACGAGTCT 102
Db 8472 ACCTTGTGCTTATCTGTGAGAGCGGGGACCCAAAGAGATGCGCGGCAACCTTACGAGTCT 853
Qy 1028 TCACGAGGCTATGACTAGGTACTCTGCCCTCGGGGACCGCGCCCAACCGGAATACG 108
Db 8532 TCACGAGGCTATGACTAGGTACTCTGCCCTCGGGGACCGCGCCCAACCGGAATACG 859
Qy 1088 ACTTGGAGCTGATAACATCGTGTCTCCAAATGCTGCGTTCGACACAGATGATCTTGGA 114
Db 8592 ACTTGGAGCTGATAACATCGTGTCTCCAAATGCTGCGTTCGACACAGATGATCTTGGA 865
Qy 1148 AAAGGGTGTACTACCTCAACCGGTGACCCCAACCGTTCCTTCCCTTGGCGGGGTGCTGGGAGA 120
Db 8652 AAAGGGTGTACTACCTCAACCGGTGACCCCAACCGTTCCTTCCCTTGGCGGGGTGCTGGGAGA 871

QY 1208 CAGCTAGGACACCTCCAGTCACTCTCGCTAGGCAACATCATCATGTATGCGCCACTT 1267
DB 8712 CAGCTAGACACACTCACTCACTCTCGCTAGGCAACATCATCATGTATGCGCCACT 8771
QY 1268 TGTGGCAAGGATGATTTCTGATGACTCACTCTCTCTCCATCTCTTAGCCAGGAGCAAC 1327
DB 8772 TATGGCAAGGATGATTTCTGATGACTCACTCTCTCTCCATCTCTTAGCCAGGAGCAAC 8831
QY 1328 TTGAAAAAGCCCTGATTTGTAGATCTAGCGGCTGTTTACTCTCATGTTACTCTC 1387
DB 8832 TTGAAAAAGCCCTGATTTGTAGATCTAGCGGCTGTTTACTCTCATGTTACTCTC 8891
QY 1388 TACCTCAGATCAATGAAGCTCCATGCTTATAGCGGCTGTTTACTCTCATGTTACTCTC 1447
DB 8892 TACCTCAGATCAATGAAGCTCCATGCTTATAGCGGCTGTTTACTCTCATGTTACTCTC 8951
QY 1448 CAGGTGAGATCAATGAGGTGGCTTATAGCGGCTGTTTACTCTCATGTTACTCTC 1507
DB 8952 CAGGTGAGATCAATGAGGTGGCTTATAGCGGCTGTTTACTCTCATGTTACTCTC 9011
QY 1508 TCTGAGACATCGGCGCAGAGTGTCCGCTCAAGTCTAGCGGCTGTTTACTCTCATGTTACTCTC 1567
DB 9012 TCTGAGACATCGGCGCAGAGTGTCCGCTCAAGTCTAGCGGCTGTTTACTCTCATGTTACTCTC 9071
QY 1568 CCATTGTGGCAAGTACCTCTTCAACTGGGCAAGTGTCCGCTCAAGTCTAGCGGCTGTTTACTCTCATGTTACTCTC 1627
DB 9072 CCATTGTGGCAAGTACCTCTTCAACTGGGCAAGTGTCCGCTCAAGTCTAGCGGCTGTTTACTCTCATGTTACTCTC 9131
QY 1628 TCCGCGCTGCTCCGCGCTGAGTGTCCGCTCAAGTGTCCGCTCAAGTCTAGCGGCTGTTTACTCTCATGTTACTCTC 1687
DB 9132 TCCGCGCTGCTCCGCGCTGAGTGTCCGCTCAAGTGTCCGCTCAAGTCTAGCGGCTGTTTACTCTCATGTTACTCTC 9191
QY 1688 ACATATATACAGAGCTGTCTCGTCCGCGCAGTGTCCGCTCAAGTGTCCGCTCAAGTCTAGCGGCTGTTTACTCTCATGTTACTCTC 1722
DB 9192 ACATATATACAGAGCTGTCTCGTCCGCGCAGTGTCCGCTCAAGTGTCCGCTCAAGTCTAGCGGCTGTTTACTCTCATGTTACTCTC 9226

RESULT 2
D89815 LOCUS D89815 9548 bp RNA linear VRL 07-MAR-1998
DEFINITION Hepatitis C virus genomic RNA, complete sequence.
ACCESSION D89815
VERSION D89815.1 GI:2943783
KEYWORDS polyprotein.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1 (sites)
AUTHORS Aizaki,H., Aoki,Y., Harada,T., Ishii,K., Suzuki,T., Nagamori,S., Toda,G., Matsura,Y. and Miyamura,T.
TITLE Full-length complementary DNA of hepatitis C virus genome from an infectious blood sample
JOURNAL Hepatology 27 (2), 621-627 (1998)
MEDLINE 98122498
PUBMED 9462666
REFERENCE 2 (bases 1 to 9548)
AUTHORS Aizaki,H.
TITLE Direct Submission
SUBMITTED (10-DEC-1996) Hideki Aizaki, National Institute of Health, Virology II; 1-23-1 Toyama, Shinjuku-ku, Tokyo, Shinjuku-ku, Tokyo 162, Japan [E-mail:aizaki@nih.go.jp, Tel:03-5285-1111, Fax:03-5285-1161]
FEATURES
source Location/Qualifiers
1..9548
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/db_xref="taxon:11103"
342..9374
/codon_start=1
/product="polyprotein"
/protein_id="BAA25076.1"

/db_xref="GI:2943784"
/translation="WSTNPKPQRTKRNTRRPODVKPPGCGQIVGGVLLPRGRPL
GVRATKTSRSEOPRRQPIPKARRRGRTRWQPGYWPWLYMGEGWGLSAPRG
SRPSWGNDRRRRLNGLVITLTCGFDLMGYIPLVGAFLGGAALAGRVRLD
VINTANLPGCSFISILLALLSLTIPASAEYEVNVSQYVHTNDSCNSIYVEAAD
GYMAHPCVPCRENSRVALPTLAARNASVPTTLRRHVDLLVGTAAFCASMY
VMDLGGVFIQLPFTFPRHETVQNCNSIYGHVSGHMDMWNNSPTAALVV
SOLLRIPOAVMDVAGAGVGLAGLAQYKSMVGNKVLVIMLFGVGDHTRVTCGVQ
GHVSTLSLPRGASOKIQLVNTGSHINRTALNCNDSUKTFLAALFYTHFNAS
GCPERMASCSIDKFDGNGPITTYADPNSDQRPYCHYAPROQGIYVPSVCGPVCY
TTFPSVVGTTDRPGATYNGDNETDVLNLRNTRPHGNWFGCTMNSTGCTKTCG
PPCINRGVGNNTLTCTDPCFKHPDATYTKGSGPMLTPRCLVDYVRLHVPCTVNF
TIFKRMVYGVGEHRLDAACNWTGRBCDRDAELSPLLLSTTEWQLPCSYTTL
PALSTGLIHUHQIVDIQYLYGIGSAVVSIALKWEYVILLPLLLADARVACALMMUL
LAQERALENLVNLNAASVYGAHMLPFNFFCAAWYMKGLRFLVGAAYFYGVMPULL
LLAALPPATMDREMVSCGGVYFGLALLTSPYCKVFLARLIMWQYFITKAEAH
LQVSLPLVNRGRDAIILLCAVHPELIIDIITKLILSLGLPMVLAQSLIRVYFVR
AQLILRACMLVRRAGGHYVQMAFLAULTGTVVVDHLTLPLOMAHVGRLDLAVP
PVFSMETKVIWGAADTAACGDIISGLPVSARRGKEILLGPADSFEQGWRLIAPIT
AVSOQVRLGLGCIITSLTGRDNQGEVGVSTAKQSFLLATCVNAGCTVFFHAGSK
TLAAAGPIQTMTINVDOLVGNPAPGARSLLTCTCGSSDLVLRHADVIYVRRRG
DSRGSLLSPRISYLGSGGGLPCPSGHVGI FRAAVCTRGVAKAVDFIPVSMETTL
MRSPFTDNTSPPAVPQTFOVHLHAPTGTGSKTKVPAAYAAQYVVLVNLVSAATL
GFGAYMSKAGHIDPNIRTVRTITTCAPITTYTGKFLADGGSGGAYDIIICDECHS
TDSTSLIGTVDLOAETVGARVWLATATPGSITTFPHNIEEVLANTGEIIPFYAK
TIPFVIRGRHLIFCHSKKCDLPAKLSALGNAYVYRGLDVSIVPASGDDVYVA
TDALMTGFTGPDSDVDCNTCTVTQVDFSLDPTFTTETTPQDASVSRQRRGTCRG
QIRYRVPVTPGERPSAMFSSVLCYDAGCAWYELTPAETS VRLRAYLNTPLGVCO
DILFEVESVFTGLTHIDAHFLSOTKQAGNFPVLVAYQATVCARAKAPPSSQNMKCC
LIRKPTLHGPTLLVRLGAVONETLTHPIIKYIMACMSADLEVTSTWLVGGVLA
ALAAVCLTGVSVIIGRIILSRPAVDPREVLYQFDEMEECASHUPYIEOGMQLAE
QPKQALGJLTQATKQAEAPVWESKRALETFWAKHMNFISGTYLAGLSLTPGN
PALASIMFTASITSLPATYTLFNLGGVAAQALPSPASAFVAGTAGAAGVCS
GIGKLVLDIAGYAGVAGALVAFKMGSDMPSTEDLVNLLPAILSPGALVGVCA
ILRRHVGPGEAVQWNNRLIAPASGNHVSPTHYVPESDAAARVTOILNLTQLK
LHOMINEDCSTPCSGWLRDWMICTVLAEFKLWLSKLLPRDGVGVFFSCQRYK
GVWRGGINYITCPCQAQITGVKNGSMRIIVGRTCSNTHGTFFPINAITYTGPSPR
APNYSRALMRVAEEYEVTRVGHVYTMGTNDNVKCPQVPAPEFFTELDGVLRR
YAPACKPLRDEVTQVGLNOYTVGSQLEPEPDVTVTSMULTDPSHITAEARRL
ARGSPSLAGSSASQLSALSKATCTTHHGAUTDILAEALRWQEMGNITRVESN
KIVLDSPFLRAERDEREVSAAELIRKTRFPAAMPVWPARDINPPLLESNVDY
VPVHGGCELPPTKAPPIPPRRKTVVLTESTVSALAEATKTFGGSGSADVSST
ATGPDQASAGDAGDSYSSMPPJEGEPDPLSDGWSVSEASEDSVCCSMS
YTWGALLTPCAEESKLPIINALSNPLLRHNMVSTTSRSASLRKOKYTFDRMOVLD
DHYRDLKEMKAKASTVAKLISVEEACKLTPHSAKSKFGKAGDVRSLSRAVNH
RSVWKLLEDYDTPIQTTIMAKNEVFCVQPEKGPKEARLIVFPDLGVRCRMALYD
VSTYUQAVMGSSYGYOYSPKORVFLVNTWKAKKCPMGFSYDFCFDSTVTEIDIR
BESIYQCDLAPARQALRSUTERYIIGPMTNSKNGCYRRCBASGVLTSCGNL
TCYLKAAACRAAKLQDCTMLVGDLLVVIDSAGTQEDAAASLRVFTAMTRYSPPG
DPPQEDLELITSCSNVSVAHDAGKRVYVLTDRDPTPLARAWEARTHTPVNSWL
GNIMYAPTLWARMLTMHTFSILIAQBLEKALDPOLYGATYSIEPLDPOIQRLLH
GLSAPLSHSYSPGEINRVASCLRLGVPLRVHRARSVRKLLSQGGRATCGKYL
FNWAVTKLKLTPPEASQLDSGFVAGYGGDIYHLSRARPRFWFMCULLLSVGV
GIYLLPNR"

ORIGIN
Query Match 91.0%; Score 1583.8; DB 14; Length 9548;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1633; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 8 GCATGTCAATGCTCTATACATGGACGCGCTGATCACCACCGCTCGCGGAGGAAA 67
DB 7594 GCTGCTCGATGCTCTACATGGACGCGCTTAATTAACACCATGCCCGCGGAGGAGA 7653
QY 68 GCAAGCTGCCCATCAACGCGCTGAGCAACTCTCTGCTGCGTCACCATAACCTGCTCTATT 127
DB 7654 GCAAGCTGCCCATTAATGCGCTGAGCAACCTCTTGTGCGCCACCAACACATGCTCTATT 7713
QY 128 CCACAACATCCCGCAGTGCAGGCTGCGGCGAGAGAGGTCACCTTTTGACAGATGCGAAG 187
DB 7714 CCACAACATCCCGCAGTGCAGGCTGCGGCGAGAGAGGTCACCTTTTGACAGATGCGAAG 7773
QY 188 TCCTGGACGATCATACCGGAGCGTGTCTAGGAGAGATGAGGCGGAGCGCTCCACAGTGA 247

VGDLOGSVFLIISQLFTFSPRHETVODNCISLYPGHVSGRNVMWMMWSPALALVV
SOLLRIPOAVDVVAGAHWGLAGLAYYSMAGNAKVLIVMLFAGVQDQIRISGQI
GHTISGLVSLFGSSQNIQLINTNGSHINRTALNCNDLNTGFLAALFYTHKNAS
GCPARMASCRSDITTAQGSIPITYAQPSISDRQPCYWHYAPRCQIVPASVCGGVYC
FPSPVAVTDRSGVTPYSGWENETDVLNNTNPRPOGNWFCCTWMSNTGFTKCGG
PTPNTGGVGNNTLTCTPDCFRKPBATYTKGSGPMLIPRCLVDYPIELMHYPCTVNF
SIFKVMYVGVGVEHRLNACNWRGRCGLDRDSELSPLLSITTEHQVLPCTSTFI
PALSTGLIHLNQNVDVQLYIGISAVSVFVWMDYVLLFLLADARACACILMMMLL
IAOAAEALNVLNNAASVAGAHGFLSFLVFFCAAYIKGRVLPQAAVALYGVNPLLL
LULALPPAVNDRMAASCGGAVLGLMLTSPYKVFYFLARLWQIYLIYTRAEAH
LQWVPLNVRGDAIIFLCTAHPBLIFDIKLLLA1LGPMLYLAGITRVPYFVR
AOGILTRACMLFKVAGHVVONAFMKLAALTGTYVDHLTPLRGHAHTGLDLAVAE
PVVSDMETKII1TWAGDIAACGDI1LGLPVSARRGBILLGPADSLEQGWRLAPIT
AISQQTGGLGCIITSLTRDNQVEGEVQVVSATQTFLLATCVNGCWTVPFHGASIK
TLAGPKPITQYTVNDQVLQVQAPPGARSMTCTCGSSDLYLVRHADYIPVRERG
DGRSLSPRPVYILKSGSGGLPCPSGHVUGIFRAAVCTRCVAKAVDPVPSNETT
MRSPVTDNSPPAVPOTFOVAHLHAPTSGSKSTKVPAAVAAQGVKVLVLPSPVAATL
GRGAYMSKAYGTDPNIRGTGRIITWGAPITYSTYKFLADGCSGGAIDIIICDCHS
TDSITLIGITVLDQAEATAGARLVLAATAPGSVTVPHNIEEVALSNTGEIPEYK
AIPETIKGRHLIFCHSKKCDLAAKLSALGINAVAYYRGLDVSIVPTSGDVVVA
TDALMTGFTGDSVIDNCTQVTFDLSLPTFTIDTTPQDAVSIVTORRGTGRG
RRGIYRVPGRVPSGMPDSSVLCBYDAGCAWVELTPAETSVRLRAYLNTPLGPVQC
DHLSPWEGVTLTRIDAHFLSQTQAGDNFPYVAYQATVCARAQAPPSWDQWKC
LTRLAPTGHPTPLIRUGAVONEVTLTHPIIKYIMTMSADLEVTVTSLVWVGVL
ALAAVCLTGSVYIVGRILISGRPAIIPDREVLRYEPDEMECEASQLEPI1EQGMQLAE
OPKALGLIQTATQAAAAPVESKWRLEAFPMGKHMNPI1SGIOYLAGLSTLPGN
PAIASMAFTAGVSTPLTOSTLNFILIGOWAAQLAPPASAAPVAGIAGAAVGS1
GIGKVLIDILAGVAGAGALVAFKVMSEMPSTEDLIVNLLPAILSPALVGVVCAA
ILRRHVGEGBAVQMNRLIAFASRGNHVSPTHVYVESDAARVQILSNLTI1QOLR
RUHOMINDSTPCSGRDVMDWICVLSDFKWLQSLKLLPRLPVFPFLSCQGRY
GVMRGDGMOTTCGCAQITGHVKNMSRI1VGPKTCSNTHGTFPI1NATYTGCTPSP
YAPSRALWRVAABEVTRVDPHYVTGMTDNVCKPCPPAPPEPTEVDEGVRLHR
YAPACKPLLRBEVTFQVGLNOYLVGSQLEPCPEPDVAVLTSMLTDPSPHITAEAKRRL
ARGSPSLASSASOLSAPSLKATCTTHDSDPADLIEANLLWRQEMGENTRVESEN
KIVLSDPLPRAEDEREISVPSILRKPFPFALPIWARPDPNPLLESWDXPDY
VYVHHGDPPLTPGKPPRRPRTVTLVSTSSALAEATKTFSGSSAADSGT
ATAPDQSDDGKSDGSYSMSPPLEPCPDPLSDGSWSVSEEGEDVWCCMS
YTWTGALI1TPCAABESKLP1INALSNLRRHNMVYATTSRACOROKKVPDRLOVLD
DHYRDLKEMKASTVAKL1SVBEACKLTPHSKSKFGYGAKDVPNLSKANVHI
RSVWKLJLEDTEPTVYMAKNEVFCVQPEKGGKPARLIVFDLGVYCEKMAVD
VSTLPQAMVSGSYQFSGPQORVEFLTNWKKCPMGFVDFTRCFDSTENDIRV
EESIYQCCDLAPARQVIRSLTERLYVGGPLTNKQNGYRRCRASGLTTSQNTL
TCYLKASAKAQLQDCTMLVNGDLDLVI1CESAGTQEDAAASLRVFTAMTRYAPPG
DPPOEYDLEI1TSCSNVSVAHADSGKRVYLLTRDPTPLABAAWETARHTPUNSWL
GNI1WYATLWARMILMTHPSI1LAQLEKALDCQIYGACYSIEPDLQI1ERLH
GLASFPLSHYSFGEINRVASCLUKLGVPLRWRHRSVRKULLSQGGRATCGKL
FNWAVRTKL1TP1PAASRLDL1SGWFVAGYSGDIYHLSRPRPRWFMLCLLLLSVGV
GIYLLPNR"

ORIGIN

Query Match	90.9%;	Score	1582.2;	DB	14;	Length	9374;
Best Local Similarity	95.2%;	Pred. No.	0;				
Matches	1632;	Conservative	0;	Mismatches	83;	Indels	0; Gaps 0;
QY	8	GCAATCTCAATGTCCTATACATGACGAGCGCCCTGATCACACCGTGCCTCGGAGGAAA	67				
Db	7582	GCTGCTCAATGTCCTATACATGACGAGCGCCCTGATCACACCGTGCCTCGGAGGAAA	7641				
QY	68	GCAAGCTGCCATCAACGCGCTGACCAACTCTCTGCTGCGTCACCATAACTGCTGTATT	127				
Db	7642	GCAAGCTGCCATCAACGCGCTGACCAACTCTTTGCTGCGTCACCAACAATGCTGTATG	7701				
QY	128	CCACAACATCCGCGAGTCAAGCCTCGCGCAGAGAAGGTCACTTTTGACAGACTGCAAG	187				
Db	7702	CTACAAACATCCGCGAGCGCATGCCAGCGCAGAGAAGGTCACTTTTGACAGACTGCAAG	7761				
QY	188	TCCTGGAGCATATTACCGGAGCTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA	247				
Db	7762	TCCTGGAGCGACACTACCGGAGCTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA	7821				
QY	248	AGGCTAAACTGTATCTGTAGAGNAGCATGCAAGCTGACCGCCCGCCGATTCGGCCAAAT	307				
Db	7822	AGGCTAAACTGTATCTGTAGAGNAGCATGCAAGCTGACCGCCCGCCGATTCGGCCAAAT	7881				

QY	308	CCAAATTTGGCTATGGGGCAAGACGCTCGGAGCCTATCCAGCAGGCGCGTTAACCCACA	367				
Db	7882	CCAAATTTGGCTATGGGGCAAGACGCTCGGAGCCTATCCAGCAGGCGCGTTAACCCACA	7941				
QY	368	TCCGCTCCGCTGTGGAAGGACTTGTGTGAGGACACTGACACACCAATTCAGACCACATCA	427				
Db	7942	TCCGCTCCGCTGTGGAAGGACTTGTGTGAGGACACTGACACACCAATTCAGACCACATCA	8001				
QY	428	TGGCAAAAAATGAGGTTTTCTCGCTCCAAACACAGAGAAGGCGCGCAAAACCGCTCGCC	487				
Db	8002	TGGCAAAAAATGAGGTTTTCTCGCTCCAAACACAGAGAAGGCGCGCAAAACCGCTCGCC	8061				
QY	488	TCATCGTATTCCTCCAGACCTCGGAGTTCTGTGTATGCGAAGATGCGCCTTTACGACGTGG	547				
Db	8062	TTATCGTATTCCTCCAGACCTCGGAGTTCTGTGTATGCGAAGATGCGCCTTTTACGACGTGG	8121				
QY	548	TTTCCACTCTTCTCCTCAGSCCGTGTATGGCTCCTCATACGGATTCATCTCTCTTAAGC	607				
Db	8122	TTTCCACTCTTCTCCTCAGSCCGTGTATGGCTCCTCATACGGATTCATCTCTCTTAAGC	8181				
QY	608	AGCGGTCGAGTTCTCTGGTGAATACCTGGAAGCAAGAAATGCGCTATGGGCTTCTCAT	667				
Db	8182	AGCGGTCGAGTTCTCTGGTGAATACCTGGAAGCAAGAAATGCGCTATGGGCTTCTCAT	8241				
QY	668	ATGACACCGCTGTTTGTGACTCAACGGTCACTGAGNATGACATCGGTGTTGAGGAGTCAA	727				
Db	8242	ATGACACCGCTGTTTGTGACTCAACGGTCACTGAGNATGACATCGGTGTTGAGGAGTCAA	8301				
QY	728	TTTACCATTTGTGACTTGGCCCCCGAAGCTAGACAGGCGCCATAAGTTCGCTCACAGAGC	787				
Db	8302	TTTATCAATTTGTGACTTGGCCCCCGAAGCTAGACAGGCTATAGGCTCGCTCACAGAGC	8361				
QY	788	GGCTCTATGTGCGGGGTCCTCAATGACTAATCTCAAAGGCGAGAACTGCGGCTATCGCGGT	847				
Db	8362	GGCTCTATGTGCGGGGTCCTCAATGACTAATCTCAAAGGCGAGAACTGCGGCTATCGCGGT	8421				
QY	848	GCGCGGAGCGGCGTCTGACGACTAGCTGCGGTATACCTCAGATGCTACTTTGAAGG	907				
Db	8422	GCGCGGAGCGGAGTACTGACGACTAGCTGCGGCAATACCTCAGATGCTACTTTGAAGG	8481				
QY	908	CCGCTGACGCTGTCGAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAATGGAGAGC	967				
Db	8482	CCTCTGCGGCTGTCGAGCTGCGAGCTCCAGGACTGCACGATGCTGTGAACCGAGAGC	8541				
QY	968	ACCTTGTGTTATCTGTGAAGCGCGGAAACCCAAAGAGGAGCGCGCAAGCTACGAGTCT	1027				
Db	8542	ACCTTGTGTTATCTGTGAGAGCGCGGAAACCCAAAGAGGATGCGCGGAGCTACGAGTCT	8601				
QY	1028	TCACGGAGGCTATGACTAGGTACTCTGCCCCCTCGGGGACCGGCCCAACCGGAATACG	1087				
Db	8602	TCACGGAGGCTATGACTAGGTACTCTGCCCCCTCGGGGACCGGCCCAACCGGAATACG	8661				
QY	1088	ACTTGGAGCTATTAACATCGTGTTCCTCCAAATGTCGCTGCGACACGATGCTCTGGCA	1147				
Db	8662	ACTTGGAGTGTATAACATCATGCTCTCCAAATGTCGCTGCGACACGATGCTCTGGCA	8721				
QY	1148	AAAGGTTGTACTACTCACCGCTGACCCACCGTCCCTTTGCGCGGGCTGCGTGGGAGA	1207				
Db	8722	AGAGGTTGTACTACTCACCGCTGACCCACCGTCCCTTTGCGCGGGCTGCGTGGGAGA	8781				
QY	1208	CAGTAGGCACTCTCCAGTCAACTCTTGGGTAGGCAACATCATCATGTATGCGGCCACTT	1267				
Db	8782	CAGTAGACACACTCTCCAGTTAACTCTCTGGCTAGGCAATATCATCATGTATGCGGCCACTT	8841				
QY	1268	TGTGGGCAAGGATGATCTGATGACTCACTCTTCTCCATCTCTTAGAGCCAGAGCAAC	1327				
Db	8842	TGTGGGCAAGGATGATCTGATGACTCACTCTTCTCCATCTCTTAGAGCCAGAGCAAC	8901				
QY	1328	TTGAAAAGCCCTGGATTGTGAGTCTACGGGCTGTGTACTCCATTGAGCCACTTCACC	1387				
Db	8902	TTGAAAAGCCCTGGATTGTGAGTCTACGGGCTGTGTACTCCATTGAGCCACTTCACC	8961				

Qy	1388	TACCTCAGATCATTAACGACTCCATGCTCTTAGCGCATTTTCACTCCATAGTTACTCTC	1447
Dd	8962	TACCTCAGATCATTAAGGAGCTCCATGCTTAGCGCATTTCACTCCATAGTTACTCTC	9021
Qy	1448	CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCACTCTGGCAG	1507
Dd	9022	CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCACTCTGGCAG	9081
Qy	1508	TCTGGAGATCGGGCCAGAGTGTCCGCGTAAGTACTGTCCAGCGGGGGAGGGCGG	1567
Dd	9082	TCTGGAGATCGGGCCAGAGTGTCCGCGTAAGTACTGTCCAGCGGGGGAGGGCGT	9141
Qy	1568	CCATTGTGGCAAGTACTCTTCAACTGGGCAGTAAGGACCAAGCTTAAACTCACTCCAA	1627
Dd	9142	CCATTGTGGCAAGTACTCTTCAACTGGGCAGTAAGGACCAAGCTTAAACTCACTCCAA	9201
Qy	1628	TTCCGGCTCGCTCCCGGCTGGACTTGTCCGCTGGTTGCTTGGCTACAGCGGGGAG	1687
Dd	9202	TCCCGCTCGCTCCCGGCTGGACTTGTCCGCTGGTTGCTTGGCTACAGCGGGGAG	9261
Qy	1688	ACATATATCAGAGCTGTCTCGTGGCGGACCCCGC	1722
Dd	9262	ACATATATCAGAGCTGTCTCGTGGCGGACCCCGC	9296
RESULT 4			
AF165059		9379 bp	RNA linear
LOCUS			Hepatitis C virus strain MD8-1 complete genome.
DEFINITION			Hepatitis C virus strain MD8-1 complete genome.
ACCESSION			AF165059
VERSION			AF165059.1
KEYWORDS			GI:5918956
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			

Qy	8	GCATGTCAATGCTCTTATACATGACAGCGGCCCTGATCACCGCTGGCTGGCGAGGAGG	67
Dd	7582	GCTGCTCAATGCTCTTACATGACAGCGGCCCTGATCACCGCTGGCGAGGAGG	7641
Qy	68	GCAAGTGGCCCATCAACGGCTGAGCAATCTCTGCTGCTGCTACCATCACTGGTCTATT	127
Dd	7642	GCAAGTGGCCCATCAACGGCTTGGCAATCTCTGCTGCTGCTACCATCACTGGTCTATT	7701
Qy	128	CCACACATCCCGAGTGGCAAGCTGGGCGAGAGAGGTTCACCTTTGACAGACTGCAAG	187
Dd	7702	CCACACATCCCGTAGCGCAAGCTTGGCGAGAGAGGTTCACCTTTGACAGACTGCAAG	7761
Qy	188	TCCTGGAGCATCTATACCGGAGCTGCTCAAGGAGATGAAGGCGGAGCGCTCCACAGTGA	247
Dd	7762	TCCTGGAGCATCTATACCGGAGCTGCTCAAGGAGATGAAGGCGGAGCGCTCCACAGTGA	7821
Qy	248	AGGCTAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCCCGCGCATTCGGCCAAAT	307
Dd	7822	AGGCTAACTCTCTATCTATAGAAGAGCTGCAAGCTGACGCCCGCGCATTCGGCCAGAT	7881
Qy	308	CCAAATTTGCTATGGGCGAAGAGAGCTCGGAGCCCTATCCAGAGCGCCGTTAAACACA	367
Dd			

ORIGIN

Query Match	90.5%;	Score 1574.2;	DB 14;	Length 9379;	
Best Local Similarity	94.9%;	Pred. No. 0;			
Matches 1627;	Conservative	0;	Mismatches 88;	Indels 0;	Gaps 0;
QY	8	GCATGTC	AAATGCTCTATACATGACAGCGCGCCCTGATCACACCGTGGCTGGCGAGGAGAA	67	
DB	7582	GCTGCTCA	ATGCTCTACATGACAGCGCGCCCTGATCACCGCTGGCTGGCGAGGAGAA	7641	
QY	68	GCAAGTGG	CCCATCAACGGCTGAGCAACTCCTTGCTGGCTACCATTAACCTCGTCTATT	127	
DB	7642	GCAAGTGG	CCCATCAACCCGTTGAGCAACTCCTTGCTGGCTACCATTAACCTCGTCTACG	7701	
QY	128	CCACACAT	CCCCGAGTGGCAAGCTCGGGCAGAGAAGGTTCACCTTTGACAGACTGCAAG	187	
DB	7702	CCACACAT	CCCCGTAGCGCAAGCTCGGGCAGAGAAGGTTCACCTTTGACAGACTGCAAG	7761	
QY	188	TCCTGGAG	CATCTATACCGGAGCTGCTCAAGGAGATGAAGGCGAAGGCGTCCACAGTGA	247	
DB	7762	TCCTGGAG	CACCTACTACCGGAGCTGCTCAAGGAGATGAAGGCGAAGGCGTCCACAGTTA	7821	
QY	248	AGGCTAA	CTGCTATCTGTAGAAGAAGCATGCAAGCTGACGCCCGCGCATTCGGCCAAAT	307	
DB	7822	AGGCTAA	CTCTCTATAGAAGAAGCTGCAAGCTGACGCCCGCGCATTCGGCCAGAT	7881	
QY	308	CCAAATTT	GCTATGGGCGAAGAGCTCGGAGCCCTATCCACGAGGCGCGTTAAACCA	367	


```
Db 9022 CAGGTGAGATCANTAGGTGGCTTTCATGCTCAGGAAACTTGGGGTTCCACCCTTTCGCGAG 9081
Qy 1508 TCTGGAGACATCGGGCCAGAGTGTCCGGGCTAAGTTACTGTCCCAAGGGGGGAGGGCGG 1567
Db 9082 TCTGGAGACATCGGGCCAGAGTGTCCGGGCTAAGTTACTGTCCCAAGGGGGGAGGGCGG 9141
Qy 1568 CCATTGTGGCAAGTACCTCTTCACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAA 1627
Db 9142 CCACATGTGGCAAGTACCTCTTCACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAA 9201
Qy 1628 TTCCGGCTCGCTCCGGCTGGACTTGTCCGGCTGGTGTCTGCTGGCTACAGCGGGGAG 1687
Db 9202 TCCGGCTCGCTCCGGCTGGACTTGTCCGGCTGGTGTCTGCTGGCTTACAGCGGGGAG 9261
Qy 1688 ACATATATCACAGCTGTCTCGTGGCCGACCCCGC 1722
Db 9262 ACATATATCACAGCTGTCTCGTGGCCGACCCCGC 9296

RESULT 6
LOCUS AB057599 1773 bp RNA linear VRL 02-APR-2003
DEFINITION Hepatitis C virus gene for RNA-dependent RNA polymerase, partial
cde, clone:pt. 2.
ACCESSION AB057599
VERSION AB057599.1 GI:29466654
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Watanabe, K. and Yoshioka, K.
REFERENCE 1 Watanabe, K. and Yoshioka, K.
AUTHORS Mutation in HCV NS5B
JOURNAL Unpublished
REFERENCE 2 Watanabe, K., Yoshioka, K., Yano, M., Ukai, K. and Ito, H.
AUTHORS Direct Submission
JOURNAL Submitted (16-MAR-2001) Kazumasa Watanabe, Nagoya University School
of Medicine, Third Department of Internal Medicine, 65 Tsuruma-cho,
Showa-ku, Nagoya, Aichi 466-8550, Japan
(E-mail: kanken@med.nagoya-u.ac.jp, Tel: 81-052-7442190,
Fax: 81-052-7442208)
FEATURES
source
1..1773
/organism="Hepatitis C virus"
/virus="Hepatitis C virus"
/mol_type="genomic RNA"
/db_xref="taxon:11103"
/clone="pt. 2"
/country="Japan"
/notes="genotype 1b"
1..1773
/gene="RdRp"
<1..>1773
/gene="RdRp"
/codon_start=1
/product="RNA-dependent RNA polymerase"
/protein_id="BA066793.1"
/db_xref="GI:29466655"
/translators="SMSYTWGALITPCAABESKLIPINALNSLRHNLVVTTSRS
AGLRQKVTDFRLQVLDHYRDLKEMKAKSTVKARLLSVBEAKLTPHSAKRF
YGAQVNLSSKATNIRSVMKLLSDPTIDTTIMAKNEVFCVQPEKGGKPARLI
VFPLGVRVCEKALYDVVSTLPOAVNGSSYGRQYSPGQRFVLTWAKKCPMGFS
YDRCPDSTVENDIRVEESYOSCDLAFARQAIRSLTERLYVGLPLTNSKQNGCY
RRCRAGVLITSCGNTLTCYLKASAAACRAKLDQCTMLVNGDDLVVICSAGTQSDAA
SLRVFTAMTRYAPGFDLPQPEYDLELITSCSNVSAHADSGKRVVLTDRPTPL
ARAAMETARITPVNSWLMGNIIMYAPTLWARMILMTHFFSILLAQELERKALDQLYGA
YSGIEPLDLPQIIRLHGLSAPLSHSYSPGIEINRVASCLRLKGLVPLRVRHRAKSVR
AKLLSGGSAATCGKLPFWAVRTKLIPFAASRLDLSGFWFVAGYSGGDIYHLSLR
ARPRFMCLLLLSVGVGIYLPNRR"
```

```
Query Match 90.4%; Score 1572.4; DB 14; Length 1773;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
Qy 13 TCAATGTCTTATACATGACAGAGGCGCCCTGATCACACCGTGGCTGCGGAGGAAGCAAG 72
Db 1 TCAATGTCTTATACATGACAGAGGCGCCCTGATCACACCGTGGCTGCGGAGGAAGCAAG 60
Qy 73 CTGGCCATCAACGGCTGAGCAACTCTCTGCTGGCTACCAATACCTGCTTATTTCCACA 132
Db 61 CTGGCCATCAACGGCTTGGAGCAACTCTCTGCTGGCTACCAATACCTGCTTATTTCCACA 120
Qy 133 ACATCCCGCAGTCAAGCTTGGCGAGAGAGTCACTTTTGACAGACTGCAAGTTCCTG 192
Db 121 ACATCCCGCAGTCAAGCTTGGCGAGAGAGTCACTTTTGACAGACTGCAAGTTCCTG 180
Qy 193 GACGATCAATACCGGAGCTGCTCAAGAGAGATGAAGCGAGGCGTCCACAGTCAAGGCT 252
Db 181 GACGATCAATACCGGAGCTGCTCAAGAGAGATGAAGCGAGGCGTCCACAGTCAAGGCT 240
Qy 253 AAATGCTATCTGTAGAGAGAGTCAAGCTGACGCGCCCGCATTCGGCCAAATCCAAA 312
Db 241 AGACTTCTATCTGTAGAGAGAGCTGCAAGCTGACGCGCCCGCATTCAGCCAAATCCAAA 300
Qy 313 TTTGGCTATGGGGCAAGGAGCTCCGAGGCTATCCAGCAGGCGCTTAACCAATCCGC 372
Db 301 TTTGGCTATGGGGCAAGGAGCTCCGAGGCTATCCAGCAGGCGCTTAACCAATCCGC 360
Qy 373 TCCGTGTGGAGAGACTTGTCTGGAGGACACTGACACCAATTCAGACACATCATCGCA 432
Db 361 TCCGTGTGGAGAGACTTGTCTGGAGGACTCTGACACCAATTCAGACACATCATCGCA 420
Qy 433 AAAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGGCGCGCAACAGCTCGCTCATC 492
Db 421 AAAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGGCGCGCAACAGCTCGCTCATC 480
Qy 493 GTATTCACAGCTGGGAGTTCGTGTATGCGAGAGAGATGCGCCCTTTACAGCTGTTTC 552
Db 481 GTATTCACAGCTGGGAGTTCGTGTATGCGAGAGAGATGCGCCCTTTACAGCTGTTTC 540
Qy 553 ACTCTTCTCAGCGCGTGTATGGCTCTCTCATACGGAATTCATCTCTCTTAAGCAGCG 612
Db 541 ACCCTTCTCAGCGCGTGTATGGCTCTCTCATACGGAATTCATCTCTCTTAAGCAGCG 600
Qy 613 GTCGAGTTCCTGCTGAATACCTGGAAGCAAGAAATGCGCTATGGCTTCTCATATGAC 672
Db 601 GTCGAGTTCCTGCTGAATACCTGGAAGCAAGAAATGCGCTATGGCTTCTCATATGAC 660
Qy 673 ACCCGCTGTTTTCAGCTCAACGGTCACTGAGAAATGACATCGCTGTCAGGAGTCAATTTAC 732
Db 661 ACCCGCTGTTTTCAGCTCAACGGTCACTGAGAAATGACATCGCTGTCAGGAGTCAATTTAC 720
Qy 733 CAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCGCAAGTTCGCTCAAGAGCGGCTC 792
Db 721 CAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCGCAAGTTCGCTCAAGAGCGGCTC 780
Qy 793 TATGTCGGGGTCCCATGACTAATTCCAAAGGCGAGAACTGCGGCTATCGCGGTCGCGC 852
Db 781 TATGTCGGGGTCCCATGACTAATTCCAAAGGCGAGAACTGCGGCTATCGCGGTCGCGC 840
Qy 853 GCGAGCGGCTGCTGACGACTAGCTCGGCTAATACCTCAGATGCTACTTTGAAGCGGCT 912
Db 841 GCGAGCGGCTGCTGACGACTAGCTCGGCTAATACCTCAGATGCTACTTTGAAGCGGCT 900
Qy 913 GCAGCTGTCGAGCTGCCAAGCTCCAGGACTCCAGATGCTCGTGAATCGAGACGACCTT 972
Db 901 GCGGCTGTCGAGCGCGGAGCTCCAGGACTCCAGATGCTCGTGAATCGAGACGACCTT 960
Qy 973 GTCGTTATCTGTGAAGCGCGGAAACCCAAAGAGGAGCGGCGCAAGCTTACAGTCTTACG 1032
Db 961 GTCGTTATCTGTGAAGCGCGGAAACCCAAAGAGGAGTCCGCGGAGCTTACAGTCTTACG 1020
Qy 1033 GAGGCTATGACTAGTACTTCTGCCCCCTGCGGAGACCGCCCGCAACCGGAATACGACTTG 1092
```

ORIGIN

Db	7587	TCAATGTCCTACACATGGACAGGTGCCCTGATCACCATTGCGCTGCGGAGGAGACAA	7646
Qy	73	CTGCGCCATCAACCGCTGAGCAACTCTTTGCTGCGTCAACCATAAAGCTGGTCTATTCCACA	132
Db	7647	CTGCGCCATCAACCGCTGAGCAACTCTTTGCTGCGTCAACCATAAAGCTGGTCTATTCCACA	7706
Qy	133	ACATCCGCACTGCAAGCTCTGGGCAAGAAAGTTCACCTTTGACAGACTGCAAGTCTCTG	192
Db	7707	ACATCCGCACTGCAAGCTCTGGGCAAGAAAGTTCACCTTTGACAGACTGCAAGTCTCTG	7766
Qy	193	GACGATCATTTACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTGAAGCT	252
Db	7767	GACGACCATACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTGAAGCT	7826
Qy	253	AAATGCTATCTGTAGAGAAAGATGCAAGCTGACGCGCCCGCATTTCCGCGCAATCCAAA	312
Db	7827	AAATGCTATCTGTAGAGAAAGCTGTAAGCTGACGCGCCCGCATTTCCGCGCAATCCAAA	7886
Qy	313	TTTGCTATGGGCAAAAGAGCTCCGAGCGCTATCCAGCAGGCGGCTTAACCAATCCGCG	372
Db	7887	TTTGCTATGGGCAAAAGAGCTCCGAGCGCTATCCGAGCGGCGGCTTAACCAATCCGCG	7946
Qy	373	TCCGCTGCGAAGGACTTCTGCTGAGAGACACTGACACACCAATTCAGACCAACATCATGCA	432
Db	7947	TCCGCTGCGAAGGACTTCTGCTGAGAGACACTGACACACCAATTCAGACCAACATCATGCA	8006
Qy	433	AAAATGAGGTTTTCTGCTGCTCAACAGAGAAAGGAGCGCCCAACAGCTCGGCTCATC	492
Db	8007	AAAATGAGGTTTTCTGCTGCTCAACAGAGAAAGGAGCGCCCAACAGCTCGGCTCATC	8066
Qy	493	GTATTTCCAGACCTGGGAGTTCTGTATGCGAGAAAGATGGCGCTTTACGAGCTGGTTCC	552
Db	8067	GTATTTCCAGACCTGGGAGTTCTGTATGCGAGAAAGATGGCGCTTTACGAGCTGGTTCC	8126
Qy	553	ACTTCTTCCAGCGCTGATGGCTCTCTCATACGATTCGATATCTCTCTTAAGCAGCGG	612
Db	8127	ACCCTTCTCAGCGCTGATGGCTCTCTCATACGATTCGATATCTCTCTTAAGCAGCGG	8186
Qy	613	GTCGAGTCTCTGCTGAATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCATATGAC	672
Db	8187	GTCGAGTCTCTGCTGAATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCATATGAC	8246
Qy	673	ACCGCTGTTTGTGACTCAACGGTCACTGAGATGACATCGTGTGAGGATCAATTTAC	732
Db	8247	ACCGCTGTTTGTGACTCAACGGTCACTGAGATGACATCGTGTGAGGATCAATTTAC	8306
Qy	733	CAATGTTGTGACTTGGCGCCCGAGCTAGACAGGCCATAAGTGCCTCACAGAGCGGCTC	792
Db	8307	CAATGTTGTGACTTGGCGCCCGAGCTAGACAGGCCATAAGTGCCTCACAGAGCGGCTC	8366
Qy	793	TATGTCGGGGTCCCATGACTAATCTCAAAGGGCAGAACTCGCGCTATCGCGGTGCGG	852
Db	8367	TATGTCGGGGTCCCATGACTAATCTCAAAGGGCAGAACTCGCGCTATCGCGGTGCGG	8426
Qy	853	GCGAGCGGCTGTGACGACTAGTGGGGTAATACCTCATGCTACTTTGAAGCGGCT	912
Db	8427	GCAAGCGGCTGTGACGACTAGTGGGGTAATACCTCATGCTACTTTGAAGCGGCT	8486
Qy	913	GCGAGCTGTGAGTGTGCAAGCTCCAGGACTGACGATGCTCGTGAATGGAGACGACTT	972
Db	8487	GCGAGCTGTGAGTGTGCAAGCTCCAGGACTGACGATGCTCGTGAATGGAGACGACTT	9546
Qy	973	GTCGTTATCTGTGAAGCGCGGAAACCCAGAGAGCGCGGAGCTTACGAGTCTTCACG	1032
Db	8547	GTCGTTATCTGTGAAGCGCGGGAACCCAGAGAGATGCGGAGGCTTACGAGCTTTCAG	9606
Qy	1033	GAGGCTATGACTAGTGTCTGCCCCCTTGGGACCCGCGCCCAACCGGATACGACTTG	1092
Db	9607	GAGGCTATGACTAGTGTCTGCCCCCTTGGGACCCGCGCCCAACCGGATACGACTTG	9666
Qy	1093	GAGCTGATACATCGTTCTCTCCATGTGTGCTGCGACAGATGCATCTTGGCAAAAGG	1152
Db	8667	GAGCTGATACATCGTTCTCTCCATGTGTGCTGCGACGATGCATCTTGGCAAAAGG	8726

Qy	1153	GTGTAATACCTACCCGTGACCCACCGTCCCTTGTGCGCGGTGCTGCGGAGACAGCT	1212
Db	8727	GTGTAATACCTACCTACCTGAGACCCACCCACCCCTTGCACCGGCTGCTGCGGAGACAGCT	8786
Qy	1213	AGGCACACTCCAGTCAACTCTGCTAGGCAACATCATCATGTATGCGCCACCTTTGCG	1272
Db	8787	AGGCACACTCCAGTCAACTCTGCTAGGCAACATCATCATGTATGCGCCACCTTTGCG	8846
Qy	1273	GCAAGGATGATTTCTGATGACTCACTTCTTCTCCATCTTCTAGCCAGGAGCACTTGA	1332
Db	8847	GCAAGGATGATTTCTGATGACTCACTTCTTCTCCATCTTCTAGCCAGGAGCACTTGA	8906
Qy	1333	AAAGCCCTGATTTCTGATGACTCACTTCTTCTCCATCTTCTAGCCAGGAGCACTTGA	1392
Db	8907	AAAGCCCTGATTTCTGATGACTCACTTCTTCTCCATCTTCTAGCCAGGAGCACTTGA	8966
Qy	1393	CAGATCAATGAACGAGCTCCATGCTTCTAGGCAATTTTCACTTCCATAGTACTCTCCAGGT	1452
Db	8967	CAGATCAATGAACGAGCTCCATGCTTCTAGGCAATTTTCACTTCCATAGTACTCTCCAGGT	9026
Qy	1453	GAGATCAATAGGCTGGCTTCTATGCTCAGGAAGCTTGGGGTACCACCTTTCGGAGTCTGG	1512
Db	9027	GAGATCAATAGGCTGGCTTCTATGCTCAGGAAGCTTGGGGTACCACCTTTCGGAGTCTGG	9086
Qy	1513	AGACATCGGCCCAAGAGTGTCCGCGCTAAGTCTTCTCCAGGGGGGAGCGCGCCAT	1572
Db	9087	AGACATCGGCCCAAGAGTGTCCGCGCTAAGTCTTCTCCAGGGGGGAGCGCGCCAT	9146
Qy	1573	TGTGGCAAGTACCTTCTCAACTGGGCGAGTAAGGACCAAGCTTAACTCACTTCCAAATTCG	1632
Db	9147	TGTGGCAAGTACCTTCTCAACTGGGCGAGTAAGGACCAAGCTTAACTCACTTCCAAATTCG	9206
Qy	1633	GCTGCGTCCCGGTGGAGCTTGTCCGCGTGGTCTGCTGGCTACAGCGGGGAGACATA	1692
Db	9207	GCTGCGTCCCGGTGGAGCTTGTCCGCGTGGTCTGCTGGCTACAGCGGGGAGACATA	9266
Qy	1693	TATCAGACCTGTCTCGTCCCGGACCCCG 1721	
Db	9267	TATCAGACCTGTCTCGTCCCGGACCCCG 9295	
RESULT 8			
LOCUS	AB049088	9616 bp	RNA linear VRL 22-AUG-2002
DEFINITION	Hepatitis C virus genomic RNA, complete genome, isolate:HCVT094.		
ACCESSION	AB049088		
VERSION	AB049088.1	GI:11559442	
KEYWORDS	Hepatitis C virus		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	1. Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K., Hatahara, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijioka, M., and Mishi, S. Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the 'progression score' revisited Hepatol. Res. 20 (2), 161-171 (2001)		
AUTHORS	2. (bases 1 to 9616)		
TITLE	Mishi, S.		
JOURNAL	Direct Submission		
PUBMED	Submitted (19-SEP-2000) Shunji Mishi, Toshiba General Hospital, Department of Medical Sciences, 6-3-22 Higashi Oh-i, Shinagawa-Ku, Tokyo 140-8522, Japan (E-mail:shunji.mishi@po.toshiba.co.jp, Tel:81-3-3764-8981, Fax:81-3-3764-8992)		
AUTHORS	11348851		
TITLE	2. (bases 1 to 9616)		
JOURNAL	Mishi, S.		
PUBMED	Direct Submission		
AUTHORS	Submitted (19-SEP-2000) Shunji Mishi, Toshiba General Hospital, Department of Medical Sciences, 6-3-22 Higashi Oh-i, Shinagawa-Ku, Tokyo 140-8522, Japan (E-mail:shunji.mishi@po.toshiba.co.jp, Tel:81-3-3764-8981, Fax:81-3-3764-8992)		
TITLE	11348851		
JOURNAL	2. (bases 1 to 9616)		
PUBMED	Mishi, S.		
AUTHORS	Direct Submission		
TITLE	Submitted (19-SEP-2000) Shunji Mishi, Toshiba General Hospital, Department of Medical Sciences, 6-3-22 Higashi Oh-i, Shinagawa-Ku, Tokyo 140-8522, Japan (E-mail:shunji.mishi@po.toshiba.co.jp, Tel:81-3-3764-8981, Fax:81-3-3764-8992)		
JOURNAL	11348851		
FEATURES	1. .9616		
source	/organism="Hepatitis C virus"		
	/mol type="genomic RNA"		
	/isolate="HCVT094"		

[illegible]

RESULT 11

HPCRNA	HPCRNA	9456 bp	RNA	linear	VRL 03-FEB-1999
LOCUS	Hepatitis C virus RNA, complete genome sequence.				
DEFINITION	D10934.1 GI:471116				
ACCESSION	C; E; NS1/E2; NS2; NS3; NS4; NS5; polyprotein precursor.				
VERSION	Hepatitis C virus				
KEYWORDS	Hepatitis C virus				
SOURCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.				
ORGANISM	1 (bases 1 to 9456)				
REFERENCE	Wang, Y., Okamoto, H., Tsuruta, F., Nagayama, R., Tao, Q.M. and Mishihiro, S.				
AUTHORS	Prevalence, genotypes, and an isolate (HC-C2) of hepatitis C virus				
TITLE	in Chinese patients with liver disease				
JOURNAL	J. Med. Virol. 40 (3), 254-260 (1993)				
MEDLINE	93359897				
PUBMED	8394876				
COMMENT	These data kindly submitted in computer readable form by: Hiroaki				

FEATURES

5' UTR
CDS

/product="polyprotein precursor"
 /protein_id="BA01728.1"
 /db_xref="GI:47117"
 /translation="MSTNPKPQTKRNTNRNPDVKFPGGGQVGVGYLLPARGPL
 GVHATKTESGPRGRPKPKARKREGRAWAQGYPMPLYGNEGWMAGWLLSPRG
 SRPSGPNDRPRSNIKGVLTITSCGLFADMGVIPLUGAPLGAARALAHGVAILED
 GVNATYATLPGCSFSLFALLSCLTIPASAVEYRNAGVYVHTDNCSSNIVVEAAD
 MINHPGCVPCVRENNSRGVALTPTLAARNASVPTTIRHVLDLLVGGAAALCSAMY
 VDGCSGVLSQVLETPRHRHETVQCCNGCSIPGHITLHMAWDMNMNSPTTALV
 SQLIRIPQAWMDMAAGWAGVLAGLYAVSMGNMAKVLIVLLGAVDGVNTYVGGAA
 ARGASGITLSRFSQSIKQIVNTNGSMHINRALTNCNDSPFGTICLAALFAHFRNS
 GCPERMASCRSIDKQSGWGPITVYQGDSPDQRYCHNYPTGPIVPAVEGSGPVYC
 FTSPVAVGTDRDLGVYVYNGENETDVLNNTTRPQGHWFCTNMWNTTPTKTCGG
 PCNIGAGGNNTLTPTDCFRKHPKATRGATYKCGSGFMILTPCLVDYVYRLAHBYCTV
 TIFKMYGVGEVHRDLADGKTRCEYCALBDRUSLPSLLTSTDMQILPSCPTTL
 PALSTGLIHNRNIVDOVLYIGISAVVYAFKWEYVLLIFLLADARVCALMMMLL
 LAQEAALENLVNLAASVAGVHGILSLVAFCAAMVYIKGRVLPGEAAVYFVGMPLLL
 LLALPRAYAMAREMASCAGAPVGLALLTSPYTKVFLPRLIMWLOFYTIRAEAF
 LQWIPLEVRGGRDALILTCAHPALLPTDKLLALFEPIMVAGQYIKMYPFVR
 AQGLIRACMLRVAGGVHYQWAFKAPALLTGVYDHLIPRDAHTGLRLDAVAVE
 PVFMSMETKIIITWGAATAACGDITGLIPVSARRGKEILFGAPDSLGGCVGTLLAPIT
 AYSOQTRLLGCIITSLERGDKNQGEVGVYVSTQATSGFSLGCVGTVYHAGGS
 TLAGKGPITQMYNTWGLDQWGPASGILSPVAFCAAMVYIKGRVLPGEAAVYFVGMPLLL
 DSRGSLSPRPVSYLKGSSSGPLCPSGSHAAGITFRAACTVGRKAVADVPFVSMERT
 NMRSPFTDNSPAPVQFQTVAHALHAPTSGSKSNKVPAAVAGQYKVLNPLSVAATL
 GFAGYMSKAHGDIDNTRVTITGAPITVYITKGLFADGCSGGADYLLICECHS
 TDSTTILIGITVQDAETAGRLVLAATPQSVYVPHSNIIEVALSTLIPCYFGK
 ALPIETIKGRHILFCHSKCKDQELAAKSLAGLNNAVAYRGLDQVYPTSGDGVVVA
 TDAIMLPTGCDPDSVTDNCTVQTFSLDPTTETITVPGDAVSRSGRRGTGVR
 RRGYIRVTPGERPSGMFDSVLCEQADGACWYBLPAETYSRLVARLANTPGDLPVCC
 DHELEWSEVTPGLTHIDALFALQKQADNHPPIVAYQATVCARAQAPPTSGWGLVLA
 LIRLPTLHPTGLTHIDALFALQKQADNHPPIITKYIMACNSADLEVTSTWVLGGVLA
 ALAAAYCTGGSVIVYVLSGRPAVIDRELYVEQFDEMECSUHLPIEQGMQLAE
 QFKQALGLLQIATKQEAALAAVPVSRKRALTEYFQKWMKMFNIGSVYLAGLITPGN
 PAIASLWATASVTPSLTQSTLLFIMGGVAAQALPAGVAGVAGIAGAAGVCS
 GLGVKVLIDLAGVAGALVAFKIMSGETSAEDLVNLLPAILSPGALGVGVCA
 IQRHVGPGSAGVOMNRNLIFAFSGRNHVSPHYVPESDAARVQTLSSITITQLLK
 RLHWINEDCTSPGCSWLRIIDWMTISVQDFTKWGLSKLPAIPALPGPFSCQGYPS
 RWGMDGIMOTTPCGAQTITGVNGYSMRIGVGTCSNTHGTFPINAYTTPGCTPYS
 APNYSRALWAAVEAYREVYRGDHPYITGMTDNIKPCQVPAWPEFTYVDGVLRLH
 VAPACKVILREVDYFQVGLNQYDPVGSQLPCEBPZLMTLMDIPSHITAEAAKRLR
 ARGSPGLASSASQLSAPSLKATCTHSDPSADLISANLLIROMEGMNTIRVESEN
 KWLIDFEPURAEEDEREVSVAEILUKTRTRPPAMPVWAPDYNPIJESKWDPDV
 VVPVHGCPPLTPKAPPIPPRRKTRVLTSTVSALAAHLAKITGSGSGSADVSGT
 ATAPDOTSNDGRSDAESYMSMPELGECDPDLSDGSMSTVESBASVDWCCSM
 YTMATQITPCAAEESKLPINLSDILRHNNVATTSRASAQRKQKTFDRLQVLI
 DRYSDVILKEMKASTVKAALLSIIEACKLTPHSAKSKFGYKAGDVRNLSCKAVNH
 RSVKDWLDETETPIDTTIMAKNEVFCQVPEKRGKVPFVDPCTPOSTVTEINDIR
 VYSTLPQAWGSSYGTQSPQGVFELNTWKSCKCPMGFSYDTPCTPOSTVTEINDIR
 BESTYQCDTAAPEAKQDLSTLDELIVGGPLTNSNGONGCYRRKCRASGVLTSCGNTI
 TCYILKASACARAKQDLCTMLVNGDILWICESAGTQDAASLRVFTBAMTYSAPPT
 DPPOPEYDLELITSCSSNVFASHDGKRYVLYETROPITPTLAAAXEATHTPNSVNI
 GNIIMWATPLWMLNTHPEFSILLAQEOLKALECQIVGACYSLEPDLQIIRLI
 GLUSAPLSHASYGELINRVASCLKGLVPLPVRWHRASVRAKLISQGRRAATCGKY
 FNWATYRKLKLTIPPAASRLDLSGSWFVAGYVGGDIYHISLSRARPRFVMLCLLLLSL
 GYLLLPNR"

9375. .9456
GTYLEFNK
/note="with a T-stretch of 45 bp."

3' UTR
ORIGIN

Query Match 89.6%; Score 1558.2; DB 14; Length 9456;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 1617; Conservative 0; Mismatches 98; Indels 0;

8	GCATGTCAATGTCCTATATACATGACAGGCGCCCTGATATACACCGTGGCTCGGAGGAA	67
2y		
2b		
7594	GCTGCTCAATGTCCTACATGACAGGTGCCTAATACACATGCGCTCGGAGGAA	765
68	GCAAGCTGCCATCAACGGCGCTGAGCAACTCTTCTGCGTACCAATACCTGGTCTATT	127
2y		
2b		
7654	GCAAGCTGCCATCAACCCCTCTGAGCAACTCTTCTGCTGGCTACCAACAATGGTGTATG	771
128	CCCAACATCCCGAGTGAAGCGCTCGGCGAGAAGAAGGTCACTTTGACAGCATGCAAG	187

DPPREYDLLELITSCSSNVSVAHDAASGKVVYLTREDPTTPLAPAAWETARHTFPWNSWL
GNIIMYAPTEWARMLIMTHFFSILLAEQLEKTLDOCOIYGACYSIEBLDLPOLLERLH
GLSAPLSHSYSPGINVRKSLGVPPLVRWRHRARSVRKLLSOGSRAATCGKYL
FNWAVRTKULKTPIPAASRLDLSWFVAGYSGGDIYHLSLRARPRWFMLCILLLSGV
GIYLLPNR"

ORIGIN

Query Match		89.3%;	Score 1554;	DB 14;	Length 9379;
Best Local Similarity		94.2%;	Pred. No. 0;		
Matches 1614;		Conservative	0;	Mismatches	100;
				Indels	0;
				Gaps	0;
QY	8	GCATGTCAATGTCCTATACATGGACAGGCGCCTGATCACACCGTGGCCCTCGGGAGGAA	67		
DB	7582	GCTGTTCCGATGCTTACACATGGACAGGCGCCTGATCACGCGTGGCTCGGGAGGAA	7641		
QY	68	GCAAGCTGCCGATCAACGGCTGAGCACTCTCTGCTGCTCACCATAACTGGTCTATT	127		
DB	7642	GCAAGTTGCCATCAACCGCTGAGCAACTCTCTGCTGCTCACCATAAATATGTTCTATG	7701		
QY	128	CCACAACATCCCGCAGTCAAGCTCGGCGAGAGAGGTCACTTTTGACAGACTGCAAG	187		
DB	7702	CCACAACATCCCGCAGTGCAGGCTACGGCAGAGAGGTCACTTTGACAGACTGCAAG	7761		
QY	188	TCTGACGATCATTTACCGGGACGTCTCAAGGAGATGAAGCGGAGGCGTCCACAGTGA	247		
DB	7762	TCTGGACGACCACTACCGGGACGTCTTAAGGAGATGAAGCGGAGGCGTCCACAGTTA	7821		
QY	248	AGCTAAACTGCTATCTGTAGAAGAGCATCAAGCTGACGCGCCCGCATTCGGGCCAAAT	307		
DB	7822	AGCTAAACTCTTATCTATAGAAGAGAGCTGCAAGCTGACGCGCCCGCACATTCGGGCCAAAT	7881		
QY	308	CCAATTTTGGCTATGGGCAAGAGCGTCCGGAGCCTATCCAGCAGGCGGCTTAACCCACA	367		
DB	7882	CCAATTTTGGCTATGGGCAAGAGCGTCCGGAGCCTATCCAGCAGGCGGCTTAACCCACA	7941		
QY	368	TCCGCTCCGTTGCGAAGGACTTGTCTGGAGGACACTGACACACCAATTTAGACCCACATCA	427		
DB	7942	TCCGCTCCGTTGCGAAGGACTTGTCTGGAGGACATGAGACACCAATTTAGACCCACATCA	8001		
QY	428	TGGCAAAAATGAGGTTTTCTCGTCCAAACAGAGAAAGAGGCGCCCAACCGAGCTCGCC	487		
DB	8002	TGGCAAAAATGAGGTTTTCTCTGTGTCCAAACAGAGAAAGGAGGCGCCCAAGAGCTCGCC	8061		
QY	488	TCATCGTATTTCCAGACCTGGGAGTTTCTGTATGCGAGAGATGCGCCCTTTAGACGTGG	547		
DB	8062	TTATCGTATTTCCAGATTTTGGGGTCCGTGTATGCGAGAGATGCGCCCTTTAGACGTGG	8121		
QY	548	TTTCCACTCTTCTCAGGCGGTGATGGCTCCTCATACCGATTCCAATACTCTCTTAAGC	607		
DB	8122	TTTCCACTCTTCTCAGGCGGTGATGGGCTCCTCATACCGATTCCAATACTCTCTTAAGC	8181		
QY	608	AGCGGGTCGAGTTTCCCTGGTGAATACTGTGAAGCAAGAAATGCCCTATGGGTTCTCAT	667		
DB	8182	AGCGGGTCGAGTTTCCCTAGTGAATGCTTTGGAATCAAAGAAATGCCCTATGGGCTTTTCA	8241		
QY	668	ATGACACCGCTGTTTGAATCAACGGTCACTGAGATGATCATCCGTGTTGAGAGTCAA	727		
DB	8242	ATGACACCGCTGTTTGAATCAACGGTCACTGAGAGTGACATCCCGTGTGAGAGTCAA	8301		
QY	728	TTTACCAATGTTGACTTTGGCCCCCGAAGCTGACAGGCGCATTAAGTTCGCTCACAGAGC	787		
DB	8302	TCTACCAATGTTGACTTTGGCCCCCGAAGCGACAGGCGCATTAAGTTCGCTCACAGAGC	8361		
QY	788	GGCTCTATGTCGGGGTCCCATGACTAACTCAAAGGCGAGAACTGCGGCTATCGCCGGT	847		
DB	8362	GGCTCTATATCGGGGTCCTGACCAATTCAAATTCAAAGGCGAGAACTGCGGTTATCGCCGGT	8421		
QY	848	CGCGCGGAGCGGGTGTGACGACTAGCTCGGTAATACCCCTCAATGCTACTTTGAAGG	907		
DB	8422	CGCGCGGAGCGGGTGTGACGACTAGCTCGGTAATACCCCTTACATGTTTACTTTGAAGG	8481		
QY	908	CCGCTGCAGCTGTGAGCTGCCAAGCTCCAGGACTCCACGATGCTCGTGAATCGAGCG	967		

Db	8482	CCTCTGAGCCTGTGAGCTGCAAAAGCTCCAGGACTGCAAGATGCTCGTGAACGGAGACG	8541
QY	968	ACCTTGTGCTTACTGTGAAAGCGCGGAAACCAAGAGGACCGCGGAAGCTTACGAGTCT	1027
Db	8542	ACCTTGTGCTTACTGTGAGAGCGCGGAAACCAAGAGGACCGCGGAAGCTTACGAGTCT	8601
QY	1028	TCACGGAGGCTATGACTTAGGTACTCTGTCCCGCCCTCTGGGAGCCCGCCCAACCGGATAACG	1087
Db	8602	TCACGGAGGCTATGACTTAGGTACTCTGTCCCGCCCGCGGAGACCCCGCCAGAGATAACG	8661
QY	1088	ACTTGGAGCTGATAACATTCGTTCCTCAATGTGTGCGTGGCAGACGATGCAATCTGGCA	1147
Db	8662	ACTTGGAGCTGATAACATTCGTTCCTCAACGTTGTGCGTGGCAGACGATGCAATCTGGCA	8721
QY	1148	AAAGGTGTACTACTCACCCTGACCCACCGTCCCGCTTGGCGGGCTGCGTGGGAGA	1207
Db	8722	AAAGGTGTACTACTCACCCTGACCCACCGTCCCGCTTGGACGGGCTGCGTGGGAGA	8781
QY	1208	CAGTAGGCACACTCCAGTCAACTCTCTGGCTAGGCAACATCATCATATGATGCGGCCACTT	1267
Db	8782	CAGCTAGGCACACTCCAGTCAACTCTCTGGCTAGGCAACATCATCATATGATGCGGCCACTT	8841
QY	1268	TGTGGGCAAGGATGATTCTGATGACTCACTTCTTCTCCATCTTCTTAGCCAGGAGCAAC	1327
Db	8842	TGTGGGCGAGATGATTCTGATGACTCACTTCTTCTCCATCTTCTTAGCCAGGAGCAAC	8901
QY	1328	TTGAAAAGCCCTGGATGTGCAGATCTACCGGGCTTGTACTCCATTTGAGCCAGGAGCAAC	1387
Db	8902	TTGAAAAGCCCTGGATGTGCAGATCTACCGGGCTTGTACTCCATTTGAGCCAGGAGCAAC	8961
QY	1388	TACCTCAGATCAATTAAGAGCTCCATCGTCTTACGCGATTTCTACTCCATAGTTACTCTC	1447
Db	8962	TACCTCAGATCAATTAAGAGCTCCATCGTCTTACGCGATTTCTACTCCATAGTTACTCTC	9021
QY	1448	CAGGTGAGATCAATAGGTGGCTTCTATGCTCTCAGGAAGCTTGGGGTACCACCTTGGGAG	1507
Db	9022	CAGGTGAGATCAATAGGTGGCTTCTATGCTCTCAGGAAGCTTGGGGTACCACCTTGGGAG	9081
QY	1508	TCTGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGGGGCGG	156
Db	9082	TCTGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGGGGCGG	914
QY	1568	CCATTTGTGGCAAGTACCTCTTCAACTGGGCAAGTAAAGCAAGCTTAACTCACTCCAA	162
Db	9142	CCATTTGTGGCAAGTACCTCTTCAACTGGGCGGTAAAGCAAGCTTAACTCACTCCAA	920
QY	1628	TTCCGGCTCGGTCGCGGCTGGACTTGTCCGGTGGTTCGTGTGCTGGCTACAGCGGGGAG	168
Db	9202	TCCCGAGCGGTCGCGGTTGGACTTGTCCAGCTGTTTGTGCTGCTACAGCGGGGAG	926
QY	1688	ACATATATCACAGCCTGTCTGTCGCCGACCCCG	1721
Db	9262	ACATTTATCACAGCCTGTCTGTCGCCGACCCCG	9295

RESULT 14					
AB057601	Hepatitis C virus gene for RNA-dependent RNA polymerase, partial cds, clone:pt. 4.	1773 bp	RNA	linear	VRL 02-APR-200
LOCUS	AB057601				
DEFINITION	AB057601.1	GI:29466658			
ACCESSION	AB057601				
VERSION					
KEYWORDS	Hepatitis C virus				
SOURCE	Hepatitis C virus				
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.				
REFERENCE	1				
AUTHORS	Watanabe, K. and Yoshioka, K.				
TITLE	Mutation in HCV NS5B				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1773)				
AUTHORS	Watanabe, K., Yoshioka, K., Yano, M., Ukai, K. and Ito, H.				

Direct Submission
Submitted (16-MAR-2001) Kazumasa Watanabe, Nagoya University School
of Medicine, Third Department of Internal Medicine; 65 Tsuruma-cho,
Showa-ku, Nagoya, Aichi 466-8550, Japan
(E-mail: kanken@med.nagoya-u.ac.jp, Tel: 81-052-74421190,
Fax: 81-052-7442001)

FEATURES	SOURCE
----------	--------

```

source
1. .1773
/organism="Hepatitis C virus"
/virus
/mol_type="genomic RNA"
/db_xref="taxon:1103"
/clone="pt. 4"
/country="Japan"
/note="genotype 1b"
1. .1773
/gene="RdRp"
<1. .>173
/gene="RdRp"
/codon_start=1
/product="RNA-dependent RNA p
/protein_id="BAC66795.1"
/db_xref="GI:29466659"
/translation="MSYITWTGALITPCA
ASLQKKVTPDLQVLDHYDLVKENKA
YGAQDVNLSRAVNHVIRVWEDLLEDS
VFPLDGVYVCEKMALYDVVSTLPOAVMS
YDTRCSVSTESDIRVEESIYQCDDLAP
RRCRASGVLTSCGNTLTCLKATAACRA
SLRAFTAMTAPPGDPPOPEYDULBI
LRAAEVARTHTPVNSMGLNIIWYAPTNAL
CYSLEPDLDPQIIRLHGSLSAFSLHYSY
AKLLSQGGRAATCKGYLFNWAARVTKJLI
ARPPWFVGLIIISVGVCIIYLLPNR"

```

ORIGIN

Query Match	89.3%	Score 1553.8	DB 14	Length 1773	
Best Local Similarity	94.3%	Pred. No. 0			
Matches 1612	Conservative	0	Mismatches	97	Indels
				0	Gaps
				0	
Qy	13	TCATGTCTTATACATGACAGGCGCCTGATCATCACGTCGTCGGAGGAAAGCAAG	72		
Db	1	TCAATGTCTTACATGACAGGCGCCTGATCGCATCGCATCGCTCGGAGGAAGCAAG	50		
Qy	73	CTGCGCCATCAACGCGCTGAGCAACTCTTGTGTCGTACCAATAACCTGGTCTATTTCACA	132		
Db	61	CTGCGCCATCAATCCGTTGAGCAACTCTTGTGCGCCACCACAACATGGTCTATGCTACA	120		
Qy	133	ACATCCCGCAGTGAACGCTCGGCAGAGAGGTACCTTTGACAGACTGCAAGTCCGT	192		
Db	121	ACATCCCGCAGCGCAAGCCTACGGCAGAAGAAGGTCACTTTGACAGACTGCAAGTCCGT	180		
Qy	193	GACGATCATTTACCGGGACGTCTCAAGGAGATGAAGGGAAGGCGTCCACAGTGAAGGT	252		
Db	181	GACGACCACATCCGGGACGTGCTTTAGGAGATGAAGGCGAAGGCGTCCACAGTTAAAGCT	240		
Qy	253	AAACTGCTATCTGTAGAAGAAGCATGAAGCTGACGCCGCCCGCATTCGGGCCAATCCAAA	312		
Db	241	AAACTTCTATCTATGAGGAAGCTTGCAAACTGACGCCGCCACACTCGGCCAAATCCAAG	300		
Qy	313	TTTGGCTATGGGGCAAGGACGTCGGAGGCTATCCAGCAGGGCGGTACACACATCCGC	372		
Db	301	TTTGGCTATGGGGCAAGGACGTCGGAACTATCCAGCAGGGCGGTAAACCAATCCGC	360		
Qy	373	TCCGTGTGGAAGGACTTGTCTGGAGGACACTGACACACCAATTCAGACCAACCATCATGGCA	432		
Db	361	TCCGTGTGGGAGACTTGTCTGGAAGACTCTGAGACCAATTTGACACCAACCATCATGGCA	420		
Qy	433	AAAAATAGGTTTTCGCGTCCAAACAGAGAGGCGCGCAAAACGAGCTCGCCTCATC	492		
Db	421	AAAAATAGGTTTTCGCGTCCAAACAGAGAGGCGCGCAAAACGAGCTCGCCTCATC	480		
Qy	493	GTATTCCGACCTGGGAGTTTCGTGTATGCGAAGATGGCCCTTTTACGACGTGGTTCC	552		

Db 1561 TGTGCAAGTACCTCTTCAACTGGGACGTAAGGACCAAGCTCAAACTCACTCCAATCCCG 1620
Qy 1633 GCTGCGTCCGGCTGGACTTTCGGCTGGTCTGCTGCTACAGCGGGGAGACATA 1692
Db 1621 GCTGCGTCCGGCTGGACTTTCGGCTGGTCTGCTGCTACAGCGGGGAGACATA 1680
Qy 1693 TATCACAGCTGTCTCGTCCGACCCCG 1721
Db 1681 TATCACAGCTGTCTCGTCCGACCCCG 1709

RESULT 15
AX686788 LOCUS 1743 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 2 from Patent EP1065213.
ACCESSION AX686788
VERSION AX686788.1 GI:29372348
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.

ORGANISM
1
Ago,H.; Miyano,M. and Adachi,T.
Hcv polymerase suitable for crystal structure analysis and method
for using the enzyme
Patent: EP 1065213-A 2 03-JAN-2001;
Japan Tobacco Inc. (JP)
Location/Qualifiers
1. .1743
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="DNA encoding fusion protein consisting of a portion
of HCV polymerase and histidine tag at the C-terminus"
1. .1740
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD86587.1"
/db_xref="GI:29372349"
/translation="MSMSYTWGALITPCAABESKLPINALSNSLLRHNNVYATSR
SAGLRQKVTDFRLQVDDHYRDLVKEMKAKASTVKAKLLSVEEACKLTPPHSAKSF
GYAKDVRNLSKAVNHIHSVWKDLEDVTPIDTTMAKNEVFCVQPEKGGKPEARL
IIVFDLGRVCEKMALYDVVSTLPQVVMGSSYGFQYSPQORVEFLVNTWKSXKNPMGF
SYDTRFDSVTENDIRVESIYQCDDLAPEARQAIKSILTERLYIGGLPNTSKGNQCG
YRRCASGVLTTCGNTLTCYLKASAAKRAKLQDCITMLVNGDDLVVLCESAGTQEDA
ASLRFTEAMTRYSPAPPQPPQPEYDLELITSCSNVSVAHDASGRVYLLTRDFTTP
LAAAMETARHTPVNSWLGNIIMYAPTLWARMILMTHFSLILAQBLEKALDQIYV
ACYSIEPLDLPQIERLHGLSAPLSHSYSPGEINRVASCLRLKLGVPPLVRWHRARS
RABLLSQGSEATCGKVLFNWAVTKLKLTPIPAASQLDLSGMFVAGYSGGDIYHSL
RARPGRSHHHHH"

ORIGIN
Query Match 89.3%; Score 1553.4; DB 6; Length 1743;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1620; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
Qy 10 ATGTCAATGTCCTATACATGACAGCGGCCCTTGATCACACCGTGGCTGGGAGGAAAGC 69
Db 1 ATGTCAATGTCCTATACATGACAGCGGCCCTTGATCACCGCATCGCTGGGAGGAAAGC 60
Qy 70 AAGTGCCTCATCAACGGCTGAGCAACTCTTGTGCTGGTCAACCTAATCACTGGTCTATTCC 129
Db 61 AAGTGCCTCATCAACGGCTGAGCAACTCTTGTGCTGGCCACCAATCAATGGTTATGCC 120
Qy 130 ACAACATCCCGCAGTGCNAGCCTGGCGCAGAAAGGTCACCTTTTGACAGACTGCAAGTC 189
Db 121 ACAACATCTCCGACGGCGGCTTGGCGCAGAAAGGTCACCTTTTGACAGACTGCAAGTC 180
Qy 190 CTGGACGATCATTTACCGGCGAGTGTCTCAAGGAGATGAAGCGGAAGGGGTCCACAGTGAAG 249
Db 181 CTGGACGACCACTACCGGCGAGTGTCTCAAGGAGATGAAGCGGAAGGGGTCCACAGTGAAG 240

Qy 250 GCTAAACTGCTATCTCTGTAAGAAGCATGCAAGCTGACCGCCCCCGCATTTGGGCCAAATCC 309
Db 241 GCTAAACTCTCTATCCGCTAGAGAAAGCTGCAAGCTGACCGCCCCCACATTTGGGCCAAATCC 300
Qy 310 AAATTTGGCTATGGGCAAGAGCGTCCGAGCGCTATCCAGCAGAGGCGCGTTAACCAACATC 369
Db 301 AAGTTTGGCTATGGGCAAGAGCGTCCGGAACCTATCCAGCAGAGGCGCGTTAACCAACATC 360
Qy 370 CGCTCCGCTGGAAGGACTTGTCTGAGGACACTGACACACCAATTCAGACCAACCAATCATG 429
Db 361 CACTCCGCTGGAAGGACTTGTCTGGAAGACATGTGACACCAATTCAGACCAACCAATCATG 420
Qy 430 GCAAAAATCAGGTTTTTCTGCTCCAAACAGAAAGGAGGCGCGCAACACAGCTCGGCTC 489
Db 421 GCAAAAATCAGGTTTTTCTGCTCCAAACAGAAAGGAGGCGCGTAAGCGCGGCTT 480
Qy 490 ATCGTATTCACAGACTCGGAGTTCGTGTATCGGAGAGATGGCCCTTTACGACGCTGGTT 549
Db 481 ATCGTATTCACAGACTCGGAGTTCGTGTATCGGAGAGATGGCCCTTCTATGATGTGGTC 540
Qy 550 TCGACTCTTCTCAGGCGCTGATGGGCTCCTCATACGGATTCCAAATACTCTCTCAAGCAG 609
Db 541 TCCACCCCTCTCAGGCTCGTGTATCGGAGTTCGTGTATCGGAGAGATGGCCCTTCTATGATGTGGTC 540
Qy 610 CGGCTCGAGTTCTCGTGAATACCTGGAAAGCAAGAAATGCCCTATGCGCTTCTCATAT 669
Db 601 CGAGTCAGTTCTCGTGAATACCTGGAAATCAAGAAACCAACCCATGGGCTTTTCATAT 660
Qy 670 GACACCCGCTGTTTTGACTCAACGGTCACTGAGAAATGACATCCGCTGTGTAGGAGTCAATT 729
Db 661 GACACTCGCTTTTCGACTCAACGGTCAACGAGAACGACATCCGCTGTGTAGGAGTCAATT 720
Qy 730 TACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCCATTAAGTTCGCTCAGAGCGG 789
Db 721 TACCAATGTTGTGACTTGGCCCCCGAAGCCAGACAGGCGCAATAAATCGCTCAGAGCGG 780
Qy 790 CTCTATGTGGGGGTCCCATCACTAACTCCAAAGGGGAGAACTCGCGCTATCGCGGTGC 849
Db 781 CTTTATATCGGGGTCTCTCTGACTTAATTCAAAGGGGAGAACTCGCGGTATTCGCGGTGC 840
Qy 850 CGCGAGCGCGGTCTGACGACTAGCTGCGGTAATACCTTCACATGCTACTTTGAAGGCC 909
Db 841 CGCGAGCGCGGTCTGACGACTAGCTGCGGTAACACCTTCACATGTTACTTTGAAGGCC 900
Qy 910 GCTGACAGCTGTGAGCTGCGAAGCTCCAGACTCCAGATGCTCGTGAATGGAGACAC 969
Db 901 TCTGACAGCTGTGAGCTGCGAAGCTCCAGACTCCAGATGCTCGTGAACGGAGACAC 960
Qy 970 CTTGTGTTATCTGAAAGCGGGGACCCAAAGAGGACGGGCAAGCTACGAGCTCTTC 102
Db 961 CTCGTGTTATCTGAAAGCGGGGACCCAAAGAGGACGGGCAAGCTACGAGCTCTTC 102
Qy 1030 ACGGAGGCTATGACTAGTACTCTCGCCCCCTGCGGGACCCCGCCCAACCGGAATACGAC 108
Db 1021 ACGGAGGCTATGACTAGTACTCTCGCCCCCTGCGGGACCCCGCCCAACCAAGATACGAC 108
Qy 1090 TTGGAGCTGATAAATCATCGTGTTCCTCCAAATGTGTGCTGGACACGATGCATCTGGCAAA 114
Db 1081 TTGGAGCTGATAAATCATCATGTTCTTCCAAATGTGTGCTGGTCCGCCACGATGCATCGGCAAA 114
Qy 1150 AGGGTGTACTACTCAACCGTGAACCCGCTCCCTTGGCGGGGCTCGGTGGGAGACA 120
Db 1141 AGGGTGTACTACTCAACCGTGAACCCGCTCCCTTGGCGGGGCTCGGTGGGAGACA 120
Qy 1210 GCTAGGCACACTCCAGTCAACTCTGGCTAGGCAACATCATCATGTATGCGGCCCACTTTG 126
Db 1201 GCTAGCACACTCCAGTTAACTCTGGCTAGGCAACATTTATTTATGATGCGGCCCACTTTG 126
Qy 1270 TGGGCAAGGATGATTTCTGATGACTCACTTCTTCTCCATCTCTTAGGCCAGGCAACTTT 132
Db 1261 TGGGCAAGGATGATTTCTGATGACTCACTTCTTCTCCATCTCTTAGGCCAGGCAACTTT 132
Qy 1330 GAAAAAGCCCTCGATTGTTCAGATCTTACGGGGCTTGTGTTATCTCCATTGAGGCCACTTGACCTTA 138

```

|||||
Db 1321 GAAAAAGCCCTGGACTGCAGATCTACGGGGCTGTACTCCATTGAGCCACTTGACCTA 1380
QY 1390 CCTCAGATCATTTGAAGACTCCATGGCTTTAGCGCATTTTCACTCCATAGTTACTCTCCA 1449
Db 1381 CCTCAGATCATTTGAAGACTCCATGGCTTTAGCGCATTTTCACTCCATAGTTACTCTCCA 1440
QY 1450 GGTGAGATCAATAGGGTGGCTTCATGCCCTCAGGAAGCTTGGGGTACCACCCCTTGGAGTC 1509
Db 1441 GGTGAGATCAATAGGGTGGCTTCATGCCCTCAGGAAGCTTGGGGTACCACCCCTTGGAGTC 1500
QY 1510 TGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGGCCGCC 1569
Db 1501 TGGAGACATCGGGCCAGAGCGTCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCCGCC 1560
QY 1570 ATTTGTGGCAAGTACTCTTCAACTGGGCAGTAAAGCAAGCTTAAACTCACTCCAATT 1629
Db 1561 ACTTGTGGCAAGTACTCTTCAACTGGGCAGTAAAGCAAGCTTAAACTCACTCCAATC 1620
QY 1630 CCGGCTGGCTCCGGCTGGACTTGTCCGGCTGGTTCTGTGGCTACAGCGGGGAGAC 1689
Db 1621 CCGGCTGGCTCCGGCTGGACTTGTCCGGCTGGTTCTGTGGTTACAGCGGGGAGAC 1680
QY 1690 ATATATCACGCTGTCTCGTGCCCGACCCCGCCATCACCATCAC 1740
Db 1681 ATATATCACGCTGTCTCGTGCCCGACCCCGCGATCCCATCACCATCAC 1731

```

Search completed: October 3, 2005, 18:55:33
Job time : 7694.28 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 10:33:35 ; Search time 5935.06 Seconds
(without alignments)
11159.430 Million cell updates/sec

Title: US-10-712-479-3
Perfect score: 1740
Sequence: 1 atggtagcatgcatgctc.....gccatcaccatcaccatcac 1740

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_his:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.4	2.6	891	AL516365	AL516365 AL516365
2	40.8	2.3	873	BQ938432	BQ938432 AGENCOURT
3	40.8	2.3	6636	BC043482	BC043482 Mus muscu
4	40.6	2.3	593	CA390175	CA390175 cs107d09.
5	40.6	2.3	665	7	CA423579
6	39	2.2	487	5	BX282027
7	39	2.2	685	6	CD579994
8	39	2.2	820	4	BG818891
9	39	2.2	851	7	CK192072
10	39	2.2	882	5	BQ988223
11	39	2.2	938	5	BQ720502
12	39	2.2	941	5	BQ720548
13	39	2.2	954	5	BQ673878
14	39	2.2	993	2	BZ623117
15	39	2.2	1606	3	CF620795
16	39	2.2	1737	3	CF591176
17	39	2.2	1739	3	CF605251
18	39	2.2	1740	3	CF615736
19	39	2.2	1742	3	CF605475
20	39	2.2	1747	3	CF598826
21	39	2.2	1753	3	CF624274
22	39	2.2	1758	3	CF609048
23	39	2.2	1763	3	CF624218
24	39	2.2	1769	3	CF596536

25	39	2.2	1812	3	CR609614	CR609614 full-leng
26	39	2.2	1828	3	CR610243	CR610243 full-leng
27	38.6	2.2	604	6	CD216287	CD216287 ppp2n.pk0
28	38.4	2.2	288	2	BB388175	BB388175
29	38	2.2	403	8	AQ936657	AQ936657 HSJ41-954
30	38	2.2	509	1	AI657258	AI657258 486092F08
31	38	2.2	554	8	CC094862	CC094862 CSU-K34.1
32	37.8	2.2	612	2	BE497710	BE497710 WHE955.80
33	37.8	2.2	772	9	CNS02XEP	AL218266 Tetraodon
34	37.8	2.2	869	9	CNS03W08	AL263969 Tetraodon
35	37.8	2.2	1012	9	CNS015JK	AL165305 Tetraodon
36	37.6	2.2	700	5	BU448257	BU448257 603767337
37	37.4	2.1	259	8	AZ865259	AZ865259 2M0175H18
38	37.4	2.1	546	2	AW255326	AW255326 ML335 pep
39	37.4	2.1	905	9	AG042203	AG042203 Pan trogl
40	37.2	2.1	841	1	AL516952	AL516952 AL516952
41	37.2	2.1	958	4	B1828644	B1828644 603078553
42	37.2	2.1	1007	5	BQ672629	BQ672629 AGENCOURT
43	37	2.1	274	9	CE434162	CE434162 tigr-gss-
44	37	2.1	703	6	CB687271	CB687271 CEST-13-H
45	37	2.1	879	9	CNS0276Y	AL184291 Tetraodon

ALIGNMENTS

RESULT 1
AL516365
LOCUS AL516365 891 bp mRNA linear EST 23-MAR-2004
DEFINITION AL516365 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
ACCESSION CSODA005YJ06 5-PRIME, mRNA sequence.
VERSION AL516365
KEYWORDS AL516365.3 GI:45653002
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 891)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:30490924.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6432.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODA005YJ06&c=6432.r.

FEATURES
source

1..891
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODA005YJ06"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with NotI and cloned into
the NotI and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 2.6%; Score 45.4; DB 1; Length 891;
Best Local Similarity 47.8%; Pred. No. 0.2; Indels 0; Gaps 0;
Matches 118; Conservative 3; Mismatches 126;


```

QY 37 GCCCTGATCACACCGTGGCTGGGAGGAAAGCAAGCTGCCCATCAACGCGCTGAGCAAC 96
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 582 GCGGCGCTCAACACAGAGTTGGAGGAGGAGCGGCAAGGNGANGAGCGGGAGACCAAC 641
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 97 TCCCTTGGCTGACCATTAACCTGGTCTATTCCACACATCCCGCAGTGCAGAGCTCGGG 156
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 642 CTCGTGCTCTTCGCAAGAGAGCTGAGACATGCACTCTGTCCCGCTGGAATAGAGGC 701
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 CAGAAAGAGGTTCACCTTTGACAGACTGCAAGTCTCGGACGATCAATACCGGAGCTGCTC 216
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 702 AGATGTAGTCTCTGATGATGAGATTGATTCTCTCAAGAGCTGCACGAGGAGAGCTG 761
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 AAGGAGATGAAGGCGAAGCGTCCACAGTGAAGGCTAAATGCTATCTGTAGAGAGCA 276
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 762 CGAGAMCTGCAGTGAAGTGTGAGAGCCAGCAGTGCAGCAGGTGGAGGTGGAGCCACG 821
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 TCGAAGC 283
: ||||
Db 822 GKGAAGC 828

```

```

RESULT 2
BQ938432 873 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8924837 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6467897
DEFINITION 5', mRNA sequence.

```

```

ACCESSION BQ938432
VERSION BQ938432.1 GI:22353910
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 873)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1394 row: i column: 18
High quality sequence stop: 709.
Location/Qualifiers
1..873
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6467897"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

```

FEATURES
source
1..873
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6467897"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

ORIGIN

```

Query Match 2.3% Score 40.8; DB 5; Length 873;
Best Local Similarity 48.3%; Pred. No. 4.5;
Matches 114; Conservative 0; Mismatches 122; Indels 0; Gaps 0;
QY 309 CMAATTTGGCTATCGGCGCAAGGACGTCGCGAGCCTATCCAGCGGCCCGCTTAACCAAT 368
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 192 CCATCTTGGCAGAGTTCCAGGAGACGCTCCGCAATATCATCAGTCTTTGCGCTCTT 251

```

```

QY 369 CCCTCCGTGTGAAGGACTTGCTGAGGAGCACTGACACACCAATTCAGACCACCATCAT 428
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 252 CCAGATCATGAGGCTGGTCAAGCTTCTGATAGGTGAGGGATCCGACACATGCTCTG 311
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 429 GGGAAAAATGAGGTTTTCGCTCCACAGAGAAGGAGCGCGCAAAACCGAGTCGCT 488
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 312 GACATTCATCAAGTCTTTCAGGCTTTCGCTATGTCGCAATGTCATGATGATAT 371
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 489 CATCGTATTCAGAGCTGGAGTTCGTATCGGAGAGATGCGCTTTTACGACG 544
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 372 CTTCACTATGCACTCATGCGATGCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 427
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 3
BQ043482 6636 bp mRNA linear HTC 19-NOV-2003
LOCUS Mus musculus calcium channel, voltage-dependent, alpha 1F subunit,
DEFINITION mRNA (cDNA clone IMAGE:5369391), containing frame-shift errors.
ACCESSION BQ043482
VERSION BQ043482.1 GI:28175783
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 6636)
Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, W.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Lu, X., Gibbs, R.A.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
Strausberg, R.
2 (bases 1 to 6636)
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgi.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smal, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

```

```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 6636)
Strausberg, R.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgi.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smal, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

```

REMARK

COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 86 Row: 0 Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9624972
This clone has the following problem: frame shifted.

FEATURES

source

Location/Qualifiers

1. .6636
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5369391"
/tissue_type="Eye, retina, mouse strain C57Bl/6"
/clone_lib="NIH MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 2.3%; Score 40.8; DB 3; Length 6636;
Best Local Similarity 48.3%; Pred. No. 8.4;
Matches 114; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 309 CAATTTTCGCTATGGGCAAGCAGCTCCGGAGCCTATCCAGCAGGGCCGTTAACCAT 368
Db |||||
QY 4435 CATCTTGGCGAGAGTTGAGAGGAGCTCCGCATATCTATCAGTTCTTTCGCTCT 4494
Db |||||
QY 369 CGCTCCGCTGTGGGAGGACTTCTCGGAGCACTGACACACCAATTCAGACCACCATCAT 428
Db |||||
QY 4495 CGAGTCATGAGCTGGTCAAGCTTCTGAGTAGGAGTGAGGGATCCGCACACTGCTCTG 4554
Db |||||
QY 429 GCAAAAATAGGTTTTCGCTCCACACGAGAGAGGAGGCGCCAAACAGCTCGCT 488
Db |||||
QY 4555 GACATTCATCAAGTCTTCCAGGCTTTCGCTATGTGCACTTCTCATAGCAATGATAT 4614
Db |||||
QY 489 CATCGATTCCAGACCTGGAGTTCGTGTATGCGAGAAAGATGCGCCCTTTACGACG 544
Db |||||
QY 4615 CTTTCATCATGCACTGTCATGTCATGTCATGTTTGGCAAGTGCTCTTCAGGAGC 4670
Db |||||

RESULT 4

CA390175

LOCUS

DEFINITION

ca107d09.y1 Human Retinal pigment epithelium/choroid cDNA

(Un-normalized, unamplified): cs Homo sapiens cDNA clone cs107d09

5', mRNA sequence.

CA390175

CA390175.1 GI:24721013

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 593)

Wistow, G., Bernstein, S.L., Wyatt, M.K., Farris, R.N., Behal, A.,

Touchman, J.W., Bouffard, G., Smith, D., and Peterson, K.

Expressed sequence tag analysis of human RPE/choroid for the

NIH Bank Project: Over 6000 non-redundant transcripts, novel genes

and splice variants

Mol. Vis. 8 (4), 205-220 (2002)

22103460

12107410

CONTACT: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 107 row: d column: 09

Seq primer: M13RPI reverse primer (ABI).

Location/Qualifiers

1. .593

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cs107d09"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retinal pigment epithelium/choroid cDNA"
(Un-normalized, unamplified): cs

/note="Organ: Eye; Vector: pCMVSPORT6; Two different donor eyes (75-80 years old) yielded approximately 600 mg of dissected RPE/choroid tissue. This in turn yielded 340 ug of total RNA and 7 ug of mRNA. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies (Rockville, MD; now part of Invitrogen Corp), essentially following the protocols of the SuperScript Plasmid System (Invitrogen Corp).
<<http://www.invitrogen.com/>>. The library code designation was cs. For this library, cDNA inserts were cloned into the NotI/MluI sites of the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 2.3%; Score 40.6; DB 6; Length 593;
Best Local Similarity 50.8%; Pred. No. 4.5;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 60 GGAGGAAGCAAGCTGCCCATCAACGGCTGAGCAACTCTTGTGCTGCCGTACCATACCT 119
Db |||||
QY 12 GTACCAAAACAGGGTGCATGAACTGGCGAGACCTGCTTCTCTCCCGCAGCATCTCT 71
Db |||||
QY 120 GGTCTATTCCACAACATCCCGCAGTGCAGCTGGCGAGAAGAGTCACTTTGACAG 179
Db |||||
QY 72 CTCCACACAGCAATACGGAAGAGTACACATTCTGTAGATGAAGCTATGACCAAGAA 131
Db |||||
QY 180 ACTGCAAGTCTTGGACCATCTTACCGGAGCTGCTCAAGGAGATGAAGCGAAGCGTC 239
Db |||||
QY 132 ATGCCATGTCCTCCGTCACCAAGCATGATTTTGCACACAGATGAATTAGAAGGAG 191
Db |||||
QY 240 CACAGTGAAGG 250
Db |||||
QY 192 GAAGATGAACG 202
Db |||||

RESULT 5

CA423579

LOCUS

DEFINITION

17000531603590 GRN_EB Homo sapiens cDNA 5', mRNA sequence.

CA423579

CA423579.1 GI:47411173

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 665)

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,

Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,

Lebkowski, J. and Stanton, L.W.

Transcriptome characterization elucidates signaling networks that

control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6), 707-716 (2004)

Contact: Brandenberger R

Regenerative Medicine

Genon Corporation

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbrandenberger@genon.com

Insert Length: 665 Std Error: 0.00.

Location/Qualifiers

1. .665

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryonic bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from HES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Query Match 2.3%; Score 40.6; DB 7; Length 665;
Best Local Similarity 50.8%; Pred. No. 4.7;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 60 GGAGGAAGCAAGTGGCCATCAACGGCTGAGCAACTCTTGGCTCACCATAACCT 119
DB 350 GTACCAAAACAGGTGGCATGAACCTGGCAGACCTGCTTCTCTCCCCAGCATCT 409
QY 120 GGTCTATTCCACACATCCGGCAGTCCAGGCTGGCGGAGAGAGGTACCTTTGACAG 179
DB 410 CTCCACACAGCAGATAGCAAGAGTACACATTTCTGTAGTGAAGCTATGACCAAGAA 469
QY 180 ACTGCAAGTCTGGACCATATTACCGGACGTGCTCAAGGAGATCAAGGGGAAAGGCGTC 239
DB 470 ATGCCATGTCCCGTCCACACAGGAGTATTTGCAACAGATGAATTAGNAGGAGG 529
QY 240 CACAGTGAAGG 250
DB 530 GAAGATGAACG 540

RESULT 6
BX282027 487 bp mRNA linear EST 04-MAR-2003
LOCUS
DEFINITION
IMAGE:4914811, mRNA sequence.
BX282027
BX282027.1 GI:28612877
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 487)
Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.
Human UnigeneSet - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998E2010822.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/CloneCards/cgi-
bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13u, Primer sequence: CGTGTAAACGAGCGCCAGT.
LOCATION/Qualifiers
FEATURES
source
1. .487
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998E2010822 ; IMAGE:4914811"
/tissue_type="anaplastic oligodendroglioma with lp/19q

loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn67"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 2.2%; Score 39; DB 5; Length 487;
Best Local Similarity 47.4%; Pred. No. 12;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCCTGATCACACCGTGGCTGGGAGGAAGCAAGCTGCCATCAACGGCTGAGCAAC 96
DB 11 GCGGCGCTCAAGCAGAGGTTGGAGGAGGACCGCGCAGCGGAGGACCGGGAGCACAAC 70
QY 97 TCCCTGCTGGTCCACATAACCTGGTCTATTTCACAAACATCCCGCAGTGCAAGCGCTGCGG 156
DB 71 CTCGTGCTCTCCCAAGGACGTGGAGTGCACCTCTGTCCGCTTGGAACTAGAGCGC 130
QY 157 CAGAAGAAGGTCACTTTTGACAGACTGCAAGTCTCTGACCATCATTTACCGGAGCTCTC 216
DB 131 AAGATTGAGTCTCTGATGGATGAGATTGAGTTCTCTCAAGAAGCTGCACGAGGAGGAGCTG 190
QY 217 AAGGAGATGAAGCGAAGGCTCCACAGTGAAGCTAACTGCTATCTGTAGAAAGAGCA 276
DB 191 CGAGACCTTCAGGTGAGTGTGGAGAGCCAGCAGGTGCAGCGTGGAGGTGGAAAGCCACG 250
QY 277 TCGAAGC 283
DB 251 GTGAAGC 257

RESULT 7
CD579994 685 bp mRNA linear EST 18-MAY-2004
LOCUS
DEFINITION
EST PF001449 hEx1 (RZPD no. 800) Homo sapiens cDNA clone
MPMGp800L12569 5' similar to PERIPHERIN.
[Source:SWISSPROT;Acc:E41219], mRNA sequence.
CD579994
CD579994.2 GI:47495907
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 685)
Bussow,K., Cahill,D., Niefeld,W., Bancroft,D., Scherzinger,B.,
Lehrach,H. and Walter,G.
A method for global protein expression and antibody screening on
high-density filters of an arrayed cDNA library
Nucleic Acids Res. 26 (21), 5007-5008 (1998)
98451599
9776767
COMMENT
On Jun 13, 2003 this sequence version replaced gi:31744319.
Contact: Bussow K
PSF, http://www.proteinstrukturfabrik.de
Max Planck Institute of Molecular Genetics
Inhestr. 73, 14195 Berlin, Germany
Tel: +49 30 32639 2802
Fax: +49 30 32639 2833
Email: buessow@molgen.mpg.de
PSF clone ID = 13446. Clone expresses a soluble His-tag fusion
protein of 45 kDa size (good expression)
Seq primer: pQE65, TGAGCGATACAAATTCACAG
High quality sequence stop: 411.
LOCATION/Qualifiers
FEATURES
source
1. .685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NPMGp800L12569"

```

/tissue_type="brain"
/dev stage="embryonic, 14.8 and 15.8 weeks (determined by
foot length)"
/clone_lib="hex1 (EZPD no. 800)"
/notes="vector: pOE3ONST (AF074376); Site 1: SalI; Site 2:
NotI; 1st strand cDNA was prepared from mRNA obtained from
brain material of two feti, 14.8 and 15.8 weeks old, with
a Not I - oligo(dT) primer [5',
GACTAGTCTTAGATCCGAGCGCCGCCCTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Sal I adaptors,
digested with Not I and cloned into the Not I and Sal I
sites of the pOE3ONST vector. This cDNA expression
library represents a subset of clones that were selected
for protein expression and has been normalised by
oligonucleotide fingerprinting. The library was
constructed by Konrad Bussow."

```

ORIGIN

Query Match	2.2%	Score 39	DB 6	Length 685
Best Local Similarity	47.4%	Pred. No. 14		
Matches 117	Conservative 0	Mismatches 130	Indels 0	Gaps 0

QY	37	GCCTGATCACACCGTGGCTGGCGAGGAAGCAAGCTGCCATCAACGGCTGAGCAAC	96
Db	36	GCGCCTCAACAGAGAGTTGGAGGAGAGCGGCAAGCGGGAGGACGGGAGCACAC	95
QY	97	TCCTTGCTGGCTACCAATACCTGGTCTATTCAACAACATCCGCGAGTCAAGCTCGG	156
Db	96	CTCGTGCTCTTCGCGAAGGACGTGGAGATGCCACTCTGTCGGCTCGAACTAGAGCG	155
QY	157	CAGAAGAAGGTCAACCTTTGACAGACTGCAAGTCTCTGACCATCATATACCGGGAGTGCTC	216
Db	156	AAGATTGAGTCTCTGATGGATGAGATTGAGTTCTCTAAGAAGCTGCACGAGGAGGAGCTG	215
QY	217	AAGGAGATGAAGCGCAAGGCGTCCACAGTGAAGCTAAACTGCTATCTGTAGAGAGCA	276
Db	216	CGAGACTCGAGTGAGTGTGGAGAGCCAGAGTGCAGCGTGGAGGTGGAGGCCACG	275
QY	277	TGCAAGC	283
Db	276	GTGAAGC	282

RESULT 8
 BG818891
 LOCUS
 DEFINITION
 602779275F2 NCI CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4914811
 linear EST 22-MAY-2001
 820 bp mRNA
 5', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 820)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM10822 row: e column: 20
 High quality sequence stop: 783.

FEATURES

source	Location/Qualifiers
	1. .820
	/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon.9606"
/clone="IMAGE:491481"
/tissue_type="anaplastic oligodendroglioma with lp/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn67"
/name="Organ: brain; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

```

ORIGIN

Query Match	2.2%	Score 39;	DB 4;	Length 820;
Best Local Similarity	47.4%;	Pred. No. 15;		
Matches 117;	Conservative 0;	Mismatches 130;	Indels 0;	Gaps 0;

QY	37	GCCTGATCACCGTGCCTGCGGAGGAAGCAAGTGCCTATCAACGCGCTGAGCAAC	96
Db	11	GCGGCGCTCAAGCAGAGGTTGGAGGAGGACGCGCAAGCGGAGGACGCGGAGCAAC	70
QY	97	TCCTTCGTGGTCACCATACCTGGTCTATTCCACAACATCCCGCAGTCAAGCCTCGG	156
Db	71	CTCGTCTCTTCGCAAGACGTGGACGATGCCACTCTGTCCGCGCTGGAACTAGAGCGC	130
QY	157	CAGAAGAAGTCACTTTGCACAGCTGCAAGTCTCTGCAGCATATTACGGGAGCTGCTC	216
Db	131	AGATTGAGTCTCTGATGATGAGTTGAGTTCTTCAAGAAGTGCACGAGGAGGAGCTG	190
QY	217	AAGGAGATGAAGCGAAGCGGTCACAGTGAAGGCTAAACTGCTATCTGTAGAAAGCA	276
Db	191	CGAGACCTCGAGTGAGTGTGGAGAGCCAGAGGTGCAGCAGGTGGAGGTGGAAAGCCAG	250
QY	277	TGCAGC	283
Db	251	GTGAGC	257

RESULT 9	CK192072	851 bp	mRNA	linear	EST 01-JUL-2004
LOCUS	EST781387	BEA	Boophilus microplus	cDNA clone	BEAD742, mRNA
DEFINITION	sequence.				
ACCESSION	CK192072				
VERSION	CK192072.1	GI:49572606			
KEYWORDS	EST.				
SOURCE	Boophilus microplus (southern cattle tick)				
ORGANISM	Boophilus microplus				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;				
AUTHORS	Parasitiformes; Ixodida; Ixodoidea; Ixodidae; Boophilus.				
JOURNAL	1 (bases 1 to 851)				
TITLE	Nene, V., Quackenbush, J., George, J. and Guerrero, F.				
COMMENT	An index of genes transcribed in the tick Boophilus microplus				
	Unpublished (2004)				
	Other ESTs: EST781386				
	Contact: Vishvanath Nene				
	Parasite Genomics				
	The Institute for Genomic Research				
	9712 Medical Center Drive, Rockville, MD 20850, USA				
	Tel: 301-610-5968				
	Fax: 301-838-0208				

	FEATURES	SOURCE
1.	100% Cotton	100%
2.	100% Cotton	100%
3.	100% Cotton	100%
4.	100% Cotton	100%
5.	100% Cotton	100%
6.	100% Cotton	100%
7.	100% Cotton	100%
8.	100% Cotton	100%
9.	100% Cotton	100%
10.	100% Cotton	100%
11.	100% Cotton	100%
12.	100% Cotton	100%
13.	100% Cotton	100%
14.	100% Cotton	100%
15.	100% Cotton	100%
16.	100% Cotton	100%
17.	100% Cotton	100%
18.	100% Cotton	100%
19.	100% Cotton	100%
20.	100% Cotton	100%
21.	100% Cotton	100%
22.	100% Cotton	100%
23.	100% Cotton	100%
24.	100% Cotton	100%
25.	100% Cotton	100%
26.	100% Cotton	100%
27.	100% Cotton	100%
28.	100% Cotton	100%
29.	100% Cotton	100%
30.	100% Cotton	100%
31.	100% Cotton	100%
32.	100% Cotton	100%
33.	100% Cotton	100%
34.	100% Cotton	100%
35.	100% Cotton	100%
36.	100% Cotton	100%
37.	100% Cotton	100%
38.	100% Cotton	100%
39.	100% Cotton	100%
40.	100% Cotton	100%
41.	100% Cotton	100%
42.	100% Cotton	100%
43.	100% Cotton	100%
44.	100% Cotton	100%
45.	100% Cotton	100%
46.	100% Cotton	100%
47.	100% Cotton	100%
48.	100% Cotton	100%
49.	100% Cotton	100%
50.	100% Cotton	100%
51.	100% Cotton	100%
52.	100% Cotton	100%
53.	100% Cotton	100%
54.	100% Cotton	100%
55.	100% Cotton	100%
56.	100% Cotton	100%
57.	100% Cotton	100%
58.	100% Cotton	100%
59.	100% Cotton	100%
60.	100% Cotton	100%
61.	100% Cotton	100%
62.	100% Cotton	100%
63.	100% Cotton	100%
64.	100% Cotton	100%
65.	100% Cotton	100%
66.	100% Cotton	100%
67.	100% Cotton	100%
68.	100% Cotton	100%
69.	100% Cotton	100%
70.	100% Cotton	100%
71.	100% Cotton	100%
72.	100% Cotton	100%
73.	100% Cotton	100%
74.	100% Cotton	100%
75.	100% Cotton	100%
76.	100% Cotton	100%
77.	100% Cotton	100%
78.	100% Cotton	100%
79.	100% Cotton	100%
80.	100% Cotton	100%
81.	100% Cotton	100%
82.	100% Cotton	100%
83.	100% Cotton	100%
84.	100% Cotton	100%
85.	100% Cotton	100%
86.	100% Cotton	100%
87.	100% Cotton	100%
88.	100% Cotton	100%
89.	100% Cotton	100%
90.	100% Cotton	100%
91.	100% Cotton	100%
92.	100% Cotton	100%
93.	100% Cotton	100%
94.	100% Cotton	100%
95.	100% Cotton	100%
96.	100% Cotton	100%
97.	100% Cotton	100%
98.	100% Cotton	100%
99.	100% Cotton	100%
100.	100% Cotton	100%

```

/organism="Boophilus microplus"
/mol_type="rRNA"
/db_xref="taxon:6941"
/clone="BEAD742"
/dev stage="Adult and larvae"
/lab_host="E. coli strain DH10B-Tona"
/clone lib="BEA"
/notes="Organ: Whole ticks and dissected organs: Vector:

```

pExpress 1: A normalized cDNA library was custom built by Express Genomics. Oligo-dT primed cDNA was directionally cloned into NotI-EcoRV site of pExpress 1. Universal M13 primers were used to generate 5' and 3' EST data."

ORIGIN

```
Query Match      2.2%; Score 39; DB 7; Length 851;
Best Local Similarity 48.4%; Pred. No. 15;
Matches 108; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 1003 GAGGACGGCGCAAGCTACGAGTCTTCACGGAGGTATGACTAGTACTCTGCCGCCCT 1062
      |||||
DB 392 GAGGACCTGGAGACCTACGACTGTCTCGCGTGTGCTGCTTACCTCATCTTTGCGTTGACGAGATTGTG 451
      |||||
QY 1063 GGGGACCGCCCAACCGGAATACGACTTGGAGCTGATAACATCGTGTTCCTCCAAATGTG 1122
      |||||
DB 452 GATGAGCAGGAGTAGCGGACAGGCATTTAACTCATCTTTGCGTTGACGAGATTGTG 511
      |||||
QY 1123 TCGGTGCGCACAGCATGCATCTGGCAAAAGGGTGTACTACCTTACCCGTGACCCACCGTC 1182
      |||||
DB 512 GCCTGGGCTACCGGAGAGCGTCAACCTGGCTCAAAATCCGACCTTCGTTCAAAATGGAC 571
      |||||
QY 1183 CCCCTTGGCGGGCTGCGTGGGAGACAGTAGGCACACTCCAG 1225
      |||||
DB 572 TCGCATGAGGAGAGGTCTACCAAGGCTGTTGACACAGACACAG 614
      |||||
```

RESULT 10

```
BQ898223
LOCUS      BQ898223
DEFINITION AGENCOURT_8065553 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6212523
            5', mRNA sequence.
ACCESSION  BQ898223
VERSION    BQ898223.1 GI:22290237
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 882)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LNCM2376 row: m column: 04
            High quality sequence stop: 588.
```

FEATURES

```
source
1..882
    Location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:6212523"
    /tissue_type="epidermoid carcinoma, cell line"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NIH MGC 102"
    /note="Organ: salivary gland; Vector: pOTB7; Site 1: XhoI;
    Site 2: EcoRI; cDNA made by oligo-dT priming.
    Directionally cloned into EcoRI/XhoI sites using the
    following 5' adaptor: GCCAGAG(G). Library constructed
    by Ling Hong in the laboratory of Gerald M. Rubin
    (University of California, Berkeley) using ZAP-cDNA
    synthesis kit (Stratagene) and Superscript II RT (Life
    Technologies). Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match      2.2%; Score 39; DB 5; Length 882;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCTGATACACCGCTGCGGAGGAAAGCAAGCTGCCATCAACCGCTTGAGCAAC 96
      |||||
DB 272 GCGCGGCTCAAGCAGAGGTTGGAGAGGAGACCGCGAAGCGGAGGACCGGAGCAAC 331
      |||||
QY 97 TCCTTGTGCTCACCATAACCTGGTCTATTCCACAACATCCCGCAGTGCACGCTTCGCG 156
      |||||
DB 332 CTGCTGCTCTTCGCGAAGGACGTGGACGATCCACTCTGTCCCGCCCTGGNACTAGAGCG 391
      |||||
QY 157 CAGAAGAAGGTCACCTTTGACAGACTGCAAGTCTCTGGACGATCATTTACCGGGACGTGCTC 216
      |||||
DB 392 AGATTGAGTCTCTGATGGATGAGATTGAGTTCTCAAGAAGCTGCACGAGGAGGAGCTG 451
      |||||
QY 217 AAGGAGATGAAGCGAAGCGCTCCACAGTGAAGGCTAAACTGCTATCTGTAGAAAGCA 276
      |||||
DB 452 CGAGACCTGCAGGTGAGTGTGGAGAGCCAGCGTGCAGCAGGTGGAGGTGGAAGCCACG 511
      |||||
QY 277 TGCAGC 283
      |||||
DB 512 GTGAAGC 518
      |||||
```

RESULT 11

```
BQ720502
LOCUS      BQ720502
DEFINITION AGENCOURT_8219550 Lupski sympathetic_trunk Homo sapiens cDNA clone
            IMAGE:6188094 5', mRNA sequence.
ACCESSION  BQ720502
VERSION    BQ720502.1 GI:21859399
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 938)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LLM13583 row: c column: 07
            High quality sequence stop: 627.
```

FEATURES

```
source
1..938
    Location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:6188094"
    /sex="male"
    /tissue_type="sympathetic trunk"
    /dev_stage="adult, 16 yr"
    /lab_host="DH10B"
    /clone_lib="Lupski sympathetic trunk"
    /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
    NotI; Site 2: SalI; cDNA made by oligo-dT priming.
    Directionally cloned using the following adaptors:
    5'-TCGACCCACCGCTCGG-3' and
    5'-GACTAGTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected
    1 kb for average insert length 1.9 kb. This is a primary
    library, non-amplified. Library constructed by Life
    Technologies and donated by J. Lupski, M.D./Ph.D. (Raylo
    College of Medicine); available through Life
    Technologies."
```

ORIGIN

Query Match 2.2%; Score 39; DB 5; Length 938;
 Best Local Similarity 47.4%; Pred. No. 15;
 Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCTGATCACACCGTGCCTGCGGAGGAGCAAGCTGCCATCAACCGCTGAGCAAC 96
 |||
 Db 230 GCGGCGCTCAAGCAGAGGTTGGAGGAGGAGACGCGCAAGCGGAGGAGCAAC 289
 |||
 QY 97 TCCTTCTGCTCACCATACCTGTCTATTCCACAACATCCGCGAGTCAAGCCCTGCGG 156
 |||
 Db 290 CTGCTGCTCTTCGCGAAGACGTGGACGATGCACTCTGTCCCGCTGGAATAGAGCGC 349
 |||
 QY 157 CAGAAGAAGTCACTTTGACAGCTGCAAGTCTTCCAGAGCTAAACTGCTATCTCTAGAAAGCA 216
 |||
 Db 350 AAGATTGAGTCTCTGATGATGAGATTGAGTTCTCTCAAGAACTGCACGAGGAGGAGCTG 409
 |||
 QY 217 AAGGAGATGAAGCGGAGCGCTCCACAGTGAAGGCTAAACTGCTATCTCTAGAAAGCA 276
 |||
 Db 410 CGAGACCTCAGGTGAGTGTGAGAGCCAGGAGTGCAGCAGGTGAGGTGGAAGCCAG 469
 |||
 QY 277 TGAAGC 283
 |||
 Db 470 GTGAAGC 476

RESULT 12
 BQ720548
 LOCUS
 DEFINITION AGENCOURT_8219731 Lupski_sympathetic_trunk Homo sapiens cDNA clone
 IMAGE:6188155 5', mRNA sequence.

ACCESSION BQ720548
 VERSION BQ720548.1 GI:21859445
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium information can be
 http://image.llnl.gov
 Plate: LLAM13583 row: e column: 20
 High quality sequence stop: 555.
 Location/Qualifiers
 1..941
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6188155"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /clone_lib="Lupski_sympathetic trunk"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
 Not1; Site 2: Sal1; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGGCTCG-3' and
 5'-GACTAGTCTTAGATCGGAGCGGCGCCCT(15)-3'. Size selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor

FEATURES

source

College of Medicine); available through Life
 Technologies."

Query Match 2.2%; Score 39; DB 5; Length 941;
 Best Local Similarity 47.4%; Pred. No. 15;
 Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCCTGATCACACCGTGCCTGCGGAGGAGCAAGCTGCCATCAACCGCTGAGCAAC 96
 |||
 Db 173 GCGGCGCTCAAGCAGAGGTTGGAGGAGGAGACGCGCAAGCGGAGGAGCAAC 232
 |||
 QY 97 TCCTTCTGCTCACCATACCTGTCTATTCCACAACATCCGCGAGTCAAGCCCTGCGG 156
 |||
 Db 233 CTCGTGCTCTTCCGCAAGACGTGGACGATGCCACTCTGTCCCGCTGGAATAGAGCGC 292
 |||
 QY 157 CAGAAGAAGTCACTTTGACAGCTGCAAGTCTTCCAGAGCTAAACTGCTATCTCTAGAAAGCA 216
 |||
 Db 293 AAGATTGAGTCTCTGATGATGAGATTGAGTTCTCTCAAGAACTGCACGAGGAGGAGCTG 352
 |||
 QY 217 AAGGAGATGAAGCGGAGCGCTCCACAGTGAAGGCTAAACTGCTATCTCTAGAAAGCA 276
 |||
 Db 353 CGAGACCTCAGGTGAGTGTGAGAGCCAGGAGTGCAGCAGGTGAGGTGGAAGCCAG 412
 |||
 QY 277 TGAAGC 283
 |||
 Db 413 GTGAAGC 419

RESULT 13
 BQ673878
 LOCUS
 DEFINITION AGENCOURT_8035094 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6212691
 5', mRNA sequence.

ACCESSION BQ673878
 VERSION BQ673878.1 GI:21784724
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium information can be
 http://image.llnl.gov
 Plate: LLCM2377 row: d column: 04
 High quality sequence stop: 492.
 Location/Qualifiers
 1..954
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6212691"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_102"
 /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-cDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies). Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 2.2%; Score 39; DB 5; Length 954;
Best Local Similarity 47.4%;
Pred. No. 15;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0

Qy	37	GCCTGATCACACCGTGGCTGGGAGGAAGCAAGCTGCCCATCAACGGCGCTGAGCAAC	96
Db	272	CGCGCGCTCAAGCAGACAGGTTGGAGGAGGACGCGCAAGCGGGAGGACGGGGAGCAAC	331

DB 332 CTCGTGCTCTCTCCGAAAGGACGTGGACGATGCCACTCTGTGTCCTGGCTTGAAGACATAGAGCGC 397

QY 157 CAGAAGAAGTCACTTTTGACAGACTGCAAGTCTCTGGACGATCATTTACCGGGACGTGTC 216

DB 392 AAGATTGAGTCTCTGATGGATGAGATTGAGTCTCTCAAGAAGCTCCAGGAGGAGGAGCIG 451

QY 217 AAGGAGATGAAGCGGAAGGGGTCCACAGTGAAGGCTAACTGTCTATCTGTAGAGAAGCA 276

DB 452 CGAGACCTGCAGGTGAGTGTGGAGAGCCAGCAGGTGCAGCAGGTGGAGAGTGGAAAGCCAGCAG 511

QY 277 TGCAGC 283
|||||

DB 512 GTGAAGC 518

PGMTT 14

Accession	Version	mRNA sequence.
BE263117		
BE263117.1	GI:9136653	

ORGANISM	SOURCE	DOI
Homo sapiens (human)		
Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: ccgahs-remail@nih.gov

Plate: LLCm16 row: g column: 07
High quality sequence stop: 699.

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="TM68.1160150"
```

```

/tissue_type="neuroblastoma"
/lab_host="DH108 (phage-resistant)"
/clone_lib="NIH MGC 19"
/name="Ovarian_brain_Vector"
/notes="pOTR7. site 1: XhoI. site 2:

```

EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG (G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 09:20:55 ; Search time 959.627 Seconds
(without alignments)
10733.697 Million cell updates/sec

Title: US-10-712-479-3

Perfect score: 1740

Sequence: 1 atggctagcatgtcaatgc.....gccatcaccatcaccatcac 1740

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1740	100.0	1740	12	Ado19022 DNA encod
2	1722	99.0	1803	12	Ado19020 DNA encod
3	1555.6	89.4	1884	6	Adg19089 HCV NS5B
4	1533.4	89.3	1743	5	Aaf24256 Hepatitis
5	1551.8	89.2	9414	13	Adr82190 Hepatitis
6	1551.8	89.2	9472	2	AAQ33282
7	1547	88.9	7979	10	Adp93729
8	1547	88.9	7979	10	Adp93730
9	1547	88.9	7979	10	Adp93732
10	1547	88.9	7979	10	Adp93731
11	1547	88.9	9595	2	AAx24843
12	1547	88.9	9595	4	AAp23492
13	1547	88.9	9595	4	AAc86939
14	1547	88.9	9595	12	Ado36222
15	1547	88.9	9595	12	Ado36222 Hepatitis
16	1545.4	88.8	7989	10	Ado79396
17	1543.8	88.7	7989	10	Ado79396 Hepatitis
18	1543.8	88.7	5955	8	ACA61114
19	1541.8	88.6	9587	13	Adr82189
20	1540.4	88.5	1776	10	AA153597
					AA153596 DNA encod

21	1540.2	88.5	1778	10	AA153598
22	1539.6	88.5	9416	2	AAQ20268
23	1539.6	88.5	9416	2	AAQ21829
24	1538	88.4	9416	2	AAx59394
25	1537.4	88.4	5965	8	ACA61111
26	1537.4	88.4	37090	8	ACA61113
27	1537	88.3	1779	2	AAZ23194
28	1537	88.3	1779	2	AAZ23195
29	1537	88.3	1779	10	AA153599
30	1535.6	88.3	9436	2	AAQ63499
31	1535.4	88.2	1779	2	AAZ23197
32	1534.8	88.2	9400	13	ADR82191
33	1534.4	88.2	1779	2	AAZ23199
34	1534.2	88.2	3123	12	ADM29102
35	1534.2	88.2	7989	10	ADD93728
36	1530.6	88.0	1779	2	AAZ23198
37	1529	87.9	1779	2	AAZ23196
38	1527.8	87.8	7141	6	AAZ23333
39	1527.8	87.8	7789	6	AAZ25330
40	1527.8	87.8	7848	6	AAZ25323
41	1527.8	87.8	7987	6	AAZ25329
42	1527.8	87.8	7987	6	AAZ25321
43	1527.8	87.8	7987	6	AAZ25324
44	1527.8	87.8	7989	3	AAx98968
45	1527.8	87.8	7989	6	AAZ25322

ALIGNMENTS

RESULT 1

AD019022

ID ADO19022 standard; DNA; 1740 BP.

XX AC ADO19022;

DT 12-AUG-2004 (first entry)

XX DE ADO19022

XX DNA encoding C-terminally truncated HCV NS5B polymerase, Cdelta21 NS5B.

KW RNA polymerase activity; continuous-read assay;

KW Hepatitis C virus infection; HCV infection; polymerase; NS5B;

KW bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;

KW diabetes; ocular disorder; renal dysfunction; lymphoma;

KW lymphoproliferative disorder; metabolic disorder; arthritis;

KW sleep disorder; thyroid disorder; antiinflammatory; hepatotropic;

KW virucide; osteopathic; cytostatic; antidiabetic; ophthalmological;

KW nephrotropic; antiarthritic; gene; ds.

XX Hepatitis C virus.

OS Hepatitis C virus.

XX Key Location/Qualifiers

FT CDS 1..1740

FT /tag= a

FT /partial

FT /product= "Cdelta21 NS5B"

XX /note= "This sequence lacks a stop codon"

PN WO2004044228-A2.

XX 27-MAY-2004.

PD 13-NOV-2003; 2003WO-US036465.

XX 13-NOV-2003; 2002US-0425981P.

PR (PFIZ) PFIZER INC.

XX Yagi Y, Sheets MP, Wells PA, Shelly JA, Poorman RA, Epps DE;

PI Morgan AG;

XX WPI; 2004-420337/39.

DR P-PSDB; ADO19023.

XX Detecting RNA polymerase activity in a continuous-read manner, useful in
PT treating osteoporosis, carcinomas, cardiovascular diseases, ocular
PT disorders or arthritis, by contacting an RNA polymerase with an
PT oligonucleotide template.

PS Disclosure; SEQ ID NO 3; 46pp; English.

XX The present invention relates to a method for detecting RNA polymerase
CC activity in a continuous-read manner. The method comprises contacting an
CC RNA polymerase with an oligonucleotide template in a reaction mixture
CC comprising an assay buffer, under conditions in which the RNA polymerase
CC is active, adding a fluorescent dye capable of binding double-stranded
CC nucleic acid molecules to the reaction mixture, and measuring the
CC fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis
CC C virus (HCV) polymerase, NS5B. The method is useful for treating HCV
CC infection, bone mineral diseases like osteoporosis, carcinomas,
CC cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,
CC lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,
CC sleep disorders and thyroid disorders. The present sequence encodes C-
CC terminally truncated HCV NS5B polymerase (designated Cdelta21 NS5B).

XX Sequence 1740 BP; 404 A; 509 C; 457 G; 370 T; 0 U; 0 Other;

Query Match 100.0%; Score 1740; DB 12; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTAGCATGTCAATGCTCCATATACATGACAGAGCGCCCTGATCACACCGTGGCTCGG 60
DB 1 ATGGCTAGCATGTCAATGCTCCATATACATGACAGAGCGCCCTGATCACACCGTGGCTCGG 60
QY 61 GAGGAAGCAAGCTGCCATCAACGCGCTGACAACTCTTCTGCTGGTCAACATTAACCTG 120
DB 61 GAGGAAGCAAGCTGCCATCAACGCGCTGACAACTCTTCTGCTGGTCAACATTAACCTG 120
QY 121 GTCTATTCCAAACATCCGCGAGTGCNAGCCTGGCGAGAGAGTCACTTTTCACAGA 180
DB 121 GTCTATTCCAAACATCCGCGAGTGCNAGCCTGGCGAGAGAGTCACTTTTCACAGA 180
QY 181 CTGCAAGTCTGACCATCATTAACGCGAGTGCNAGCCTGGCGAGAGAGTCACTTTTCACAGA 240
DB 181 CTGCAAGTCTGACCATCATTAACGCGAGTGCNAGCCTGGCGAGAGAGTCACTTTTCACAGA 240
QY 241 ACAGTCAAGGCTAAACTGTCTATCTGTAGAAGAGCATGCAAGCTGACGCCCCCGCATTCG 300
DB 241 ACAGTCAAGGCTAAACTGTCTATCTGTAGAAGAGCATGCAAGCTGACGCCCCCGCATTCG 300
QY 301 GCCAAATCCAAATTTGGCTATGCGGCAAGAGAGTCCGAGGCTTATCCAGAGGGCGGTT 360
DB 301 GCCAAATCCAAATTTGGCTATGCGGCAAGAGAGTCCGAGGCTTATCCAGAGGGCGGTT 360
QY 361 AACACATCCGCTCGGTGTGGAAGGACTTGTCTGAGAGACTGACACACCAATTCAGACC 420
DB 361 AACACATCCGCTCGGTGTGGAAGGACTTGTCTGAGAGACTGACACACCAATTCAGACC 420
QY 421 ACCATCATGCAAAATGAGGTTTCTGCGTCCAAACGAGAGAAAGAGGCGCCAAACCA 480
DB 421 ACCATCATGCAAAATGAGGTTTCTGCGTCCAAACGAGAGAAAGAGGCGCCAAACCA 480
QY 481 GCTCGCTCATCGTATTCACAGACTGGGAGTTCGTGTATGCGAGAGAGTGGCCCTTTAC 540
DB 481 GCTCGCTCATCGTATTCACAGACTGGGAGTTCGTGTATGCGAGAGAGTGGCCCTTTAC 540
QY 541 GACGTGGTTTCCACTCTTCTCAGGCGGTGATGGGCTCTCATACGGATTCGAATCTCT 600
DB 541 GACGTGGTTTCCACTCTTCTCAGGCGGTGATGGGCTCTCTATACGGATTCGAATCTCT 600
QY 601 CCTAAGCAGCGGTCGAGTTCTCTGATACCTTGGAAAGCAAGAAATCCCTATGGGC 660
DB 601 CCTAAGCAGCGGTCGAGTTCTCTGATACCTTGGAAAGCAAGAAATCCCTATGGGC 660
QY 661 TTCTCATATGACACCGCTGTTTGAATCAACGCTCACTGAGATGACATCCGCTGTGAG 720

AD019020
 ID ADO19020 standard; DNA; 1803 BP.
 XX AC ADO19020;
 XX DT 12-AUG-2004 (first entry)
 XX DE DNA encoding full-length HCV NS5B polymerase, FL NS5B.
 XX KW RNA polymerase activity; continuous-read assay;
 KW Hepatitis C virus infection; HCV infection; polymerase; NS5B;
 KW bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;
 KW diabetes; ocular disorder; renal dysfunction; lymphoma;
 KW lymphoproliferative disorder; metabolic disorder; arthritis;
 KW sleep disorder; thyroid disorder; antiinflammatory; hepatotropic;
 KW virucide; osteopathic; cytostatic; antidiabetic; ophthalmological;
 KW nephrotropic; antiarthritic; gene; ds.
 XX OS Hepatitis C virus.
 XX PH Key
 XX FT CDS Location/Qualifiers
 XX FT 1..1803
 XX FT /*tag= a
 XX FT /partial
 XX FT /product= "FL NS5B"
 XX FT /note= "This sequence lacks a stop codon"
 XX PN WO2004044228-A2.
 XX PD 27-MAY-2004.
 XX PF 13-NOV-2003; 2003WO-US036465.
 XX PR 13-NOV-2002; 2002US-0425981P.
 XX PA (PFIZ) PFIZER INC.
 XX PI Yagi Y, Sheets MP, Wells PA, Shelly JA, Poorman RA, Epps DE;
 XX PI Morgan AG;
 XX WPI; 2004-420337/39.
 XX DR P-PSDB; ADO19021.
 XX PT Detecting RNA polymerase activity in a continuous-read manner, useful in
 XX PT treating osteoporosis, carcinomas, cardiovascular diseases, ocular
 XX PT disorders or arthritis, by contacting an RNA polymerase with an
 XX PT oligonucleotide template.
 XX PS Disclosure; SEQ ID NO 1; 46pp; English.
 XX CC The present invention relates to a method for detecting RNA polymerase
 XX CC activity in a continuous-read manner. The method comprises contacting an
 XX CC RNA polymerase with an oligonucleotide template in a reaction mixture
 XX CC comprising an assay buffer, under conditions in which the RNA polymerase
 XX CC is active, adding a fluorescent dye capable of binding double-stranded
 XX CC nucleic acid molecules to the reaction mixture, and measuring the
 XX CC fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis
 XX CC C virus (HCV) polymerase, NS5B. The method is useful for treating HCV
 XX CC infection, bone mineral diseases like osteoporosis, carcinomas,
 XX CC cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,
 XX CC lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,
 XX CC sleep disorders and thyroid disorders. The present sequence encodes full-
 XX CC length HCV NS5B polymerase (designated FL NS5B).
 XX SQ Sequence 1803 BP; 414 A; 529 C; 471 G; 389 T; 0 U; 0 Other;
 Query Match 99.0%; Score 1722; DB 12; Length 1803;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCTAGCATGTCATATGCTTATATGACAGAGCGCCCTGATCACACCGTGGCTGGC 60
 DB 1 ATGGCTAGCATGTCATATGCTTATATGACAGAGCGCCCTGATCACACCGTGGCTGGC 60

61 GAGAAAGCAAGTCCCATCAACGCGTGGAGCAACTCTTCTGCTGGTCAACATACCTG 120
 DB 61 GAGAAAGCAAGTCCCATCAACGCGTGGAGCAACTCTTCTGCTGGTCAACATACCTG 120
 QY 121 GTCTATTCCACAACATCCCGAGTGCAGAGCCCTGGGCGAGAGAGAGTCACTTTGACAGA 180
 DB 121 GTCTATTCCACAACATCCCGAGTGCAGAGCCCTGGGCGAGAGAGAGTCACTTTGACAGA 180
 QY 181 CTGGAAGTCTGGAGGATCATTTACCGGACGTGCTCAAGGAGATGAAGCGGAAGCGTCC 240
 DB 181 CTGGAAGTCTGGAGGATCATTTACCGGACGTGCTCAAGGAGATGAAGCGGAAGCGTCC 240
 QY 241 ACAGTGAAGGCTAACTCTATCTGTGAAGAGCATGCAAGCTGACGCCGCCGATTCG 300
 DB 241 ACAGTGAAGGCTAACTCTATCTGTGAAGAGCATGCAAGCTGACGCCGCCGATTCG 300
 QY 301 GCCAAATCCAAATTTGGCTATGGGCGCAAGCAAGTCCGAGCCCTATCCAGGAGGCGCGTT 360
 DB 301 GCCAAATCCAAATTTGGCTATGGGCGCAAGCAAGTCCGAGCCCTATCCAGGAGGCGCGTT 360
 QY 361 AACCAATCCGCTCCGTGTGAAGGACTTGTGGAGGACACTGACACACCAATTCAGACC 420
 DB 361 AACCAATCCGCTCCGTGTGAAGGACTTGTGGAGGACACTGACACACCAATTCAGACC 420
 QY 421 ACCATCATGGCAAAATAGAGTTTCTGCTCAACAGAGAAAGGAGGCGCAAAACA 480
 DB 421 ACCATCATGGCAAAATAGAGTTTCTGCTCAACAGAGAAAGGAGGCGCAAAACA 480
 QY 481 GCTGCGCTCATGCTATTCAGAGCTGGAGTTCGTGTATGCGAGAGATGGCCCTTTAC 540
 DB 481 GCTGCGCTCATGCTATTCAGAGCTGGAGTTCGTGTATGCGAGAGATGGCCCTTTAC 540
 QY 541 GACGTGGTTCCTACTCTTCAGGCGGTGATGGGCTCCTCATACGGATTCCAATACTCT 600
 DB 541 GACGTGGTTCCTACTCTTCAGGCGGTGATGGGCTCCTCATACGGATTCCAATACTCT 600
 QY 601 CCTAAGCAGCGGTGCGAGTTCCTGCTGAATACCTTGAAGCAAGAAATGCCCTATGGGC 660
 DB 601 CCTAAGCAGCGGTGCGAGTTCCTGCTGAATACCTTGAAGCAAGAAATGCCCTATGGGC 660
 QY 661 TTCTCATATGACACCCGCTGTTTGTACTCAACGCTCACTGAGAAATGATCGGTGTGAG 720
 DB 661 TTCTCATATGACACCCGCTGTTTGTACTCAACGCTCACTGAGAAATGATCGGTGTGAG 720
 QY 721 GAGTCAATTTACCAATGTTGAGTTCGCGCCCGAGCTAGACAGGCGCATAGGTCGCTC 780
 DB 721 GAGTCAATTTACCAATGTTGAGTTCGCGCCCGAGCTAGACAGGCGCATAGGTCGCTC 780
 QY 781 ACAGAGCGGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTCGGGCTAT 840
 DB 781 ACAGAGCGGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTCGGGCTAT 840
 QY 841 CGCGGTCGCGCGAGCGCGTGTGCTGACGACTAGTTCGGTAAATACCTTCACATGCTAC 900
 DB 841 CGCGGTCGCGCGAGCGCGTGTGCTGACGACTAGTTCGGTAAATACCTTCACATGCTAC 900
 QY 901 TTGAGGCGGCTGACGCTGTCAGCTGCCAAGCTCCAGGACTGACGATGCTCGTGAAT 960
 DB 901 TTGAGGCGGCTGACGCTGTCAGCTGCCAAGCTCCAGGACTGACGATGCTCGTGAAT 960
 QY 961 GGAGAGCGCTTCGTTATCTGTGAAGCGGGGAAACCCAGAGGAGCGGCGGAGCCTA 1020
 DB 961 GGAGAGCGCTTCGTTATCTGTGAAGCGGGGAAACCCAGAGGAGCGGCGGAGCCTA 1020
 QY 1021 CGAGTCTTTCACCGAGGCTATGACTAGTACTCTGCGCCCTCGGGGACCGCCGCAACCG 1080
 DB 1021 CGAGTCTTTCACCGAGGCTATGACTAGTACTCTGCGCCCTCGGGGACCGCCGCAACCG 1080
 QY 1081 GAATACGACTTGGAGCTGATTAACATCGTGTCTTCCAAATGTGTGCGTGGCAGCATGCA 1140
 DB 1081 GAATACGACTTGGAGCTGATTAACATCGTGTCTTCCAAATGTGTGCGTGGCAGCATGCA 1140

Db	584	TTATCGTATTTCCAGACCTGGGAGTTTCGTGTATGCGAGAAATATGSCCTCTACGACGTGG	643
Qy	548	TTTTCCACTCTTCTCAGGCGGTGATGGGCTCCTCATACGGATTCCAAATCTCTCCTAAGC	607
Db	644	TCTCCACCCCTTCTCAGGCGGTGATGGGCTCCTCATACGGATTCCAAATCTCTCCTCAAGC	703
Qy	608	AGCGGTGAGTTCTGTGGTGAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT	667
Db	704	AGCGGTGAGTTCTGTGGTGAATGCTCGAATCAAGAAATGCCCTATGGGCTTCTCAT	763
Qy	668	ATGACACCCGCTGTTTGTACTCAACGGTCACTGAGAAATGACATCGGTGTGAGGAGTCAA	727
Db	764	ATGACACCCGCTGTTTGTACTCAACGGTCACTGAGAGGACATCCGTGTGAGGAGTCAA	823
Qy	728	TTTTACCAATGTGTGACTTGGGCGCGGAAGCTAGACAGCCONTAAGTGTGCTACAGAGC	787
Db	824	TCTACCAATGTGTGACTTGGGCGCGGAAGCAGACAGGCTATAAAGTCGTACAGAGC	883
Qy	788	GGCTCTATCTCGGGGTCCATGACTAACTCCAAAGGCGAGAACTGGCGCTATGCCGGT	847
Db	884	GGCTCTATATCGGGGTCCCTGACCAATTTCAAAGGGCGAAGCTGGCGGTATGCCGGT	943
Qy	848	GCCGCGAGCGCGGTGCTGACCACTAGCTGGGTAAATACCTCACAATGCTACTTTGAGG	907
Db	944	GCCGCGAGCGCGGTGCTGACCACTAGCTGGGTAAATACCTCACAATGCTACTTTGAGG	1003
Qy	908	CGCTGCAAGCTGTGAGCTGCAAGCTCCAGGACTGCACGATGCTCGTAATCGAGAGC	967
Db	1004	CCCTCTCGCGCTGTGAGCTGCAAGCTCCAGGACTGCACGATGCTCGTGAACGAGAGC	1063
Qy	968	ACCTTGTGCTTATCTGTGAAGCGCGGAACCCAAAGGACGCGGCAAGCCTACGAGTCT	1027
Db	1064	ACCTTGTGCTTATCTGTGAGAGCGCGGAACCCAAAGGAGTGCAGCAACCTACGAGTCT	1123
Qy	1028	TCACGAGGCTATGACTAGGTACTCTGCGCGCGCTGGGGAACCGCCCAACCGGAATACG	1087
Db	1124	TCACGAGGCTATGACTAGGTACTCTGCGCGCGCTGGGGAACCGCCCAACCGGAATACG	1183
Qy	1088	ACTTGGAGCTGATTAACATCTGTTCCTCAATGTGTGGTCGACACGATGCACTGCGCA	1147
Db	1184	ACTTGGAGTGTNTAATCATCTCTCTCCAATGTGTGGTCGCGCACTGCACTGCGCA	1243
Qy	1148	AAAGGTTACTACTCTACCCGTGACCCACCGTCCCTTTGCGCGGCTGCGTGGGAGA	1207
Db	1244	AAAGGTTACTACTCTACCCGTGACCCACCGTCCCTTTGCGCGGCTGCGTGGGAGA	1303
Qy	1208	CAGCTAGGACACTTCCAGTCAACTCTGGCTAGGCAACATCATGATATGCGGCCACTT	1267
Db	1304	CAGCTAGACACTCCAATCAACTCTGGCTAGGCAATATCATCATGATATGCGGCCACTT	1363
Qy	1268	TGTGGCAAGGATGATTTGATGACTCACTTCTTCTCATCTCTTAGCCCGAGAGCAAC	1327
Db	1364	TATGGCAAGGATGGTTCGATGACTCATTTCTTCTCATCTCTTAGCCCGAGAGCAAC	1423
Qy	1328	TTGAAAAGCCCTGATTTGTGAGATCTAGGGGCTGTACTCCATGAGGCACTTTGACC	1387
Db	1424	TTGAAAAGCCCTGATTTGTGAGATCTAGGGGCTGTACTCCATGAGGCACTTTGACC	1483
Qy	1388	TACCTCAGATCATTTGAACGACTCCATCGTCTTAGCGCATTTTCACTCCATAGTTACTTCT	1447
Db	1484	TACCTCAGATCATTTGAACGACTCCAAGTCTTAGCGCATTTTCACTCCACAGTTACTTCT	1543
Qy	1448	CAGGTAGATCAATAGGGTGGCTTTCATGCTCAGGAAGCTTGGGGTACCAACCTTCGAG	1507
Db	1544	CAGGTAGATCAATAGGGTGGCTTTCATGCTCAGGAAGCTTGGGGTACCAACCTTCGAG	1603
Qy	1508	TCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTATCTGTCCAGGGGGGAGGGCG	1567
Db	1604	TCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTGTCTGTCTCCAGGGGGGAGGGCT	1663
Qy	1568	CCATTTGTGCAAGTACTCTTCAACTGGGCACTTAGGACCAAGCTTAACCTCACTCCAA	1627
Db	1664	CCATTTGTGCAAGTACTCTTCAACTGGGCACTTAGGACCAAGCTTAACCTCACTCCAA	1723

QY	1628	TTCCGGCTGCTGCCGGCTGACACTGTCCGGCTGGTTTCTGTTGCTTGGCTACAGCGGGGAG	1687
Db	1724	TCCCGGCTGCGTCCCGGCTGACCTTGCTCCGGCTGTTCTGTTCTGGATACAACGGGGAG	1783
QY	1688	ACATATATACAGGCTGTCTCGTCCCGACCCCG	1721
Db	1784	ACATATATACAGGCTGTCTCGTCCCGACCCCG	1817
RESULT 4			
AAF24256			
ID	AAF24256 standard; DNA; 1743 BP.		
XX	AC	AAF24256;	
XX	DT	03-APR-2001 (first entry)	
XX	DE	Hepatitis C virus NS5B polymerase coding sequence.	
XX	KW	Hepatitis C virus; HCV; NS5B polymerase; protein coordinate data;	
XX	KW	cirrhosis; hepatoma; inhibitor; ds.	
XX	OS	Hepatitis C virus.	
XX	PN	EP1065213-A2.	
XX	PD	03-JAN-2001.	
XX	PF	30-JUN-2000; 2000EP-00113955.	
XX	PR	02-JUL-1999; 99JP-00188630.	
XX	PR	07-JUL-1999; 99JP-00192488.	
XX	PA	(NISB) JAPAN TOBACCO INC.	
XX	PI	Ago H, Miyano M, Adachi T;	
XX	DR	WPI; 2001-104834/12.	
XX	DR	P-PSDB; AAB60130.	
PT	New polypeptide, derived from hepatitis C virus (HCV) polymerase NS5B,		
PT	having a HCV polymerase activity, useful for crystal structure analysis		
PT	and for rational identification of HCV polymerase inhibitors.		
XX	Example 1; Page 274-279; 295pp; English.		
XX	The present invention provides the hepatitis C virus (HCV) NS5B		
XX	polymerase protein and its structure. It also provides methods for		
XX	identifying inhibitors of the protein which can be used in the treatment		
XX	of HCV infection. HCV can lead to cirrhosis and hepatoma in its chronic		
XX	form		
XX	Sequence 1743 BP; 415 A; 516 C; 448 G; 364 T; 0 U; 0 Other;		
Query Match 89.3%; Score 1553.4; DB 5; Length 1743;			
Best Local Similarity 93.6%; Pred. No. 0;			
Matches 1620; Conservative 0; Mismatches 111; Indels 0; Gaps 0;			
QY	10	ATGTCATCTCTATACATGGACGGCGCCCTGATCACACCGTGGCTGCGGAGGAAGC	69
Db	1	ATGTCATCTCTACATATGGACGGCGCTTGATCAGCCATCGCTGCGGAGGAAGC	60
QY	70	AAGTCCCATCAACGGCTGAGCAACTCTTGTGCGTCAACATAACCTGGTCTATTCC	129
Db	61	AAGTCCCATCAACGGCTTGAGCAACTCTTGTGCGCCACCATTAACATGGTTATGCC	120
QY	130	ACAACATCCCGAGTGCAGCGCTCGGGCAGAGAAGGTACCTTTGACAGACTGCAAGTC	189
Db	121	ACAACATCTCGAGCGAGGCTCGGGCAGAGAAGGTACCTTTGACAGACTGCAAGTC	180
QY	190	CTTGGACGATCATTTACCGGAGCGTGTCTAAGGAGATGAAGCGGAGGGCTGCACAGTGAAG	249

XX
 PA
 XX
 PI
 Manoharan M, Bumrot D;
 XX
 DR
 WPI; 2004-677362/66.
 XX
 PT
 Interference RNA agent useful for treating dyslipidemias, coronary artery
 PT
 disease, diabetes, cancer or neurological disease, comprises sense
 PT
 sequence and antisense sequence which has specific modifications.
 XX
 PS
 Example 5; SEQ ID NO 6689; 378pp; English.
 XX
 The invention describes a RNA interference (iRNA) agent (I) comprising a
 CC
 sense sequence and an antisense sequence, where the sense sequences have
 CC
 one or more asymmetrical 2'-O alkyl modifications, the antisense
 CC
 sequences have one or more asymmetrical phosphorothioate modifications
 CC
 and the antisense sequence targets a human gene sequence. Also described
 CC
 are: a pharmaceutical preparation comprising (I); reducing (M1) apoB-100
 CC
 levels or glucose-6-phosphatase levels in a subject; producing (I);
 CC
 stabilizing (I), involves selecting a sequence with activity and
 CC
 introducing one or more asymmetrical modification in the sequence, where
 CC
 the modification decreases nucleic acid sensitivity while not decreasing its
 CC
 activity; a kit comprising (I) and instruction for its use; and a device
 CC
 that can be dispense or administer a composition comprising (I). (I) is
 CC
 useful for reducing apoB-100 levels or glucose-6-phosphatase levels. (M1)
 CC
 is useful for reducing apoB-100 levels or glucose-6-phosphatase levels. (M1)
 CC
 The subject is suffering from a disorder characterized by elevated or
 CC
 otherwise unwanted expression of apoB-100, elevated or otherwise unwanted
 CC
 levels of cholesterol, and/or dysregulation of lipid metabolism. The
 CC
 disorder is chosen from the HDL/LDL cholesterol imbalance,
 CC
 dyslipidemias, hypercholesterolemia, statin-resistant
 CC
 hypercholesterolemia, coronary artery disease (CAD), coronary heart
 CC
 disease (CHD) and atherosclerosis. (I) is administered to a subject to
 CC
 inhibit hepatic glucose production or for treating glucose-metabolism-
 CC
 related disorder e.g. diabetes or type-2 diabetes. (I) is useful for
 CC
 treating the diseases as mentioned above, cancer (e.g. breast, colon or
 CC
 lung cancer), neurological disease (e.g., Huntington disease or
 CC
 spinocerebellar ataxia) or viral disease (e.g., AIDS). This sequence
 CC
 represents hepatitis C virus DNA.
 XX
 SQ
 Sequence 9414 BP; 1887 A; 2825 C; 2687 G; 2015 T; 0 U; 0 Other;
 Query Match 89.2%; Score 1551.8; DB 13; Length 9414;
 Best Local Similarity 94.1%; Pred. No. 0;
 Matches 1613; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
 QY 8 GCATGTCAATGCTTATACATGACAGAGCGCCCTGATCACACCGTGGCTGCGAGGAAA 67
 DB |||||
 QY 68 GCAAGTGGCCATCAACGGCTGAGCACTCTTGTGCTGGTCAACATTAACCTGGTCTATT 127
 DB |||||
 QY 7655 GCAAGTGGCCATCAACCGCTTGAAGATTTCTTGTCTGCTACACACACATGGTCTATG 7714
 DB |||||
 QY 128 CCACAAATCCCGCATGCAAGCTGCGGCAGAGAGGTCACTTTCACAGACTGCAAG 187
 DB |||||
 QY 7715 CTACAAATCCCGCATGCAAGCTGCGGCAGAGAGGTCACTTTCACAGACTGCAAG 7774
 DB |||||
 QY 188 TCTTGACGATCATTAACCGGCAGCTGCTCAAGAGATGAAGCGGCTCCACAGTGA 247
 DB |||||
 QY 7775 TCTTGACGATCATTAACCGGCAGCTGCTTAAAGAGATGAAGCGGCTCCACAGTGA 7834
 DB |||||
 QY 248 AGGCTAAATGCTATCTGTAGAGAGCATGCAAGCTGACGCCGCCCGCATTCGGCCAAAT 307
 DB |||||
 QY 7835 AGGCTAAATGCTATCTGTAGAGAGCATGCAAGCTGACGCCGCCCGCATTCGGCCAAAT 7894
 DB |||||
 QY 308 CCAAATTTGGCTATGGGCAAGAGAGCTCCGAGGCTATCCAGAGGGCGGTAAACACA 367
 DB |||||
 QY 7895 CCAAAATTTGGCTATGGGCAAGAGAGCTCCGAGGCTATCCAGAGGGCGGTAAACACA 7954
 DB |||||
 QY 368 TCCGCTCCGCTGGAGGAGCTTGTGGAGGACACTGACACACCAATTCAGACCACATCA 427
 DB |||||

DB 7955 TCCGCTCCGTGTGGAAGGACCTGTGGAAGACACTGAAACACCAATTAGCACTACATCA 8014
 QY |||||
 QY 428 TGGCAAAAATGAGGTTTTTCTGCTCCAAACAGAGAAAGGCGCGCAAAACAGCTCGCC 487
 DB |||||
 DB 8015 TGGCAAAAATGAGGTTTTTCTGCTCCAAACAGAGAAAGGCGCGCAAAACAGCTCGCC 8074
 QY |||||
 QY 488 TCATCGTATTCACACACCTGGGAGTTCGTGTATGCGAAGATGCGCCCTTTAGGAGTGG 547
 DB |||||
 DB 8075 TTATCGTGTCCAGATCTGGGAGTTCGTGTATGCGAAGATGCGCCCTTTATGAGCTGG 8134
 QY |||||
 QY 548 TTTCCACTCTTCTCTCAGCGCGGTGCTCTCATACGGATTCCAAATACTCTCTTAAGC 607
 DB |||||
 DB 8135 TCTCCACCTTCTCTCAGCGCGGTGCTCTCATACGGATTCCAAATACTCTCTTAAGC 8194
 QY |||||
 QY 608 AGCGGTCGAGTCTCTGCTGTAATACCTCGAAAGCAAGAAATGCCCTATGGCTTCTCAT 667
 DB |||||
 DB 8195 AGCGGTCGAGTCTCTGCTGTAATACCTCGAAATCAAGAAATGCCCTATGGCTTCTCAT 8254
 QY |||||
 QY 668 ATGACACCGCTGTTTGAAGTCAACGGTCACTGAGATGACATTCGTTTGAAGAGTCAA 727
 DB |||||
 DB 8255 ATGACACCGCTGTTTGAAGTCAACGGTCACTGAGATGACATTCGTTTGAAGAGTCAA 8314
 QY |||||
 QY 728 TTTACCAATGTTGACTTGGCGGCGGAGCTAGACAGGCTAGACAGGCTAGGCTCAGAGC 787
 DB |||||
 DB 8315 TTTACCAATGTTGACTTGGCGGCGGAGCTAGACAGGCTAGGCTCAGAGC 8374
 QY |||||
 QY 788 GGCTCTATGTCGGGGTCCCACTGACTACTCAAGAGCGAGAACTGCGGCTATCCCGGT 847
 DB |||||
 DB 8375 GGCTCTATGTCGGGGTCCCACTGACTACTCAAGAGCGAGAACTGCGGCTATCCCGGT 8434
 QY |||||
 QY 848 GCCCGGAGGCGGCTGTCAGCAGTACTGCGGTAATACCTCACAATGCTACTTTGAAGG 907
 DB |||||
 DB 8435 GCCCGGAGGCGGCTGTCAGCAGTACTGCGGTAATACCTCACAATGCTACTTTGAAGG 8494
 QY |||||
 QY 908 CGCTGACGCTGTCAGCTGCCAGTCTCAGAGCTGACGATGTCAGATGCTGTAATGAGAGC 967
 DB |||||
 DB 8495 CCACTGCGGCTGTCAGCTGCCAGTCTCAGAGCTGACGATGTCAGATGCTGTAATGAGAGC 8554
 QY |||||
 QY 968 ACCTGTCGTTATCTGTGAAGCGGGAACCCAGAGGCGGCAAGCTTACAGGCT 1027
 DB |||||
 DB 8555 ACCTGTCGTTATCTGTGAAGCGGGAACCCAGAGGCGGCAAGCTTACAGGCT 8614
 QY |||||
 QY 1028 TCAGGAGGCTATGACTAGGTACTTGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 1087
 DB |||||
 DB 8615 TCAGGAGGCTATGACTAGGTACTTGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 8674
 QY |||||
 DB 1088 ACTTGGAGCTGATTAACATCATGTTCTTCCAAATGTTGGTGCACACAGATGCATCTGCA 1147
 DB |||||
 DB 8675 ACTTGGAGTTGATAACATCATGTTCTTCCAAATGTTGGTGCACACAGATGCATCTGTA 8734
 QY |||||
 QY 1148 AAAGGTTGATCTACTTACCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1207
 DB |||||
 DB 8735 AAAGGTTGATCTACTTACCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8794
 QY |||||
 QY 1208 CAGCTAGGACACTCCAGTCAACTCTGGCTAGGCAACATCATCATGTATGGCGGCGGCGG 1267
 DB |||||
 DB 8795 CAGCTAGGACACTCCAGTCAACTCTGGCTAGGCAACATCATCATGTATGGCGGCGGCGG 8854
 QY |||||
 QY 1268 TGTGGGCAAGGATGATCTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1327
 DB |||||
 DB 8855 TATGGGCAAGGATGATCTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 8914
 QY |||||
 QY 1328 TTGAAAAAGCCTCGATTTGTCAGATCTACGGGCTTGTACTTCAATTCAGGCGGCTTACC 1387
 DB |||||
 DB 8915 TTGAAAAAGCCTCGATTTGTCAGATCTACGGGCTTGTACTTCAATTCAGGCGGCTTACC 8974
 QY |||||
 QY 1388 TACCTCAGATCATTTGAAGGCTCCATGGTCTTACGGCATTTTCACTCCATAGTTACTCTC 1447
 DB |||||
 DB 8975 TACCTCAGATCATTTGAAGGCTCCATGGTCTTACGGCATTTTCACTCCATAGTTACTCTC 9034
 QY |||||
 QY 1448 CAGGTGAGATCAATAGGCTGCTTTCATGCTCAGAGCTTGGGCTACCCCTTGGAG 1507
 DB |||||
 DB 9035 CAGGAGATCAATAGGCTGCTTTCATGCTCAGAGCTTGGGCTACCCCTTGGAG 9094

	QY	1028	TCA CGAGGCTATGACTAGGTACTTGC GCCCCTGGGGACCCGCCCAACCGGAATACG	1087
	DB	8615	TCA CGAGGCTATGACTAGGTACTTGC GCCCCTGGGGACCCGCCCTCAACCGGAATACG	8674
	QY	1088	ACTTGAGGCTGATAACATGGTGTTCCTCAAATGTGTGGTGGCACACGATCATCTGGCA	1147
	DB	8675	ACTTGAGTTGATAAACAATGTTCTCCTCCAATGTGTGGTGGCACACGATCATCTGGTA	8734
	QY	1148	AAAGGGTGTACTACCTACCCTGACCCACCGTCCCCTTGGCGGGCTGCTGGGGAGA	1207
	DB	8735	AAAGGGTGTACTACCTACCCTGACCCCTTACACCCCTTTGCA CGGGCTGCTGGGGAGA	8794
	QY	1208	CAGCTAGGCACACTCCAGTCAACTCCTGGCTAGGCAACATCATCATGTATGCGGCCACTT	1267
	DB	8795	CAGCTAGACACACTCCAGTCAACTCCTGGCTAGGCAACATCATCATGTATGCGGCCACCT	8854
	QY	1268	TGTGGGAAGGATGATTCTGATGACTCACTTCTTCTCCATCTTCTAGCCCCAGGACAA C	1327
	DB	8855	TATGGGCAAGGATGATTCTGATGACTCATTTCTTCTCCATCTTCTAGCTCAGGAGCAAC	8914
	QY	1328	TGAAAAAGCCCTGGATGTTGAGATCTACAGGGGCTTGTACTCAATTAGAGCACTTGACC	1387
	DB	8915	TGAAAAAACCCCTAGATTGTGAGATCTACAGGGGCTGTTTCTCAATTGAAACCACTTGATC	8974
	QY	1388	TACCTCAGATCATTTGAACGACTCCATGTCTTTAGCGCATTTTCACTCCATAGTTACTCTC	1447
	DB	8975	TACCTCAGATCATTTGAGCGACTCCATGTCTTTAGCGCATTTTCACTCCATAGTTACTCTC	9034
	QY	1448	CAGGTGAGATCAATAGGTGGCTTCATGCTCCAGGAAGCTTGGGGTACCACCTTGGCAG	1507
	DB	9035	CAGGCGAGATCAATAGGTGGCTTCATGCTCCAGAAACTTGGGTACACCTTGGCAG	9094
	QY	1508	TCTGGAGNATCGGGCCAGAAGTGTCCGCGCTAAGTTACTGTGCCAGGGGGAGGGCCG	1567
	DB	9095	CCTGGAGACATCGGGCCAGAAGTGTCCGCGCTAAGTACTGTGCCAGGGGGAGGGCCG	9154
	QY	1568	CCATTTGTGGCAAGTACTCTTTCAACTGGGCAATGAGCAACAGCTTAAACTCACTCCAA	1627
	DB	9155	CCACTTTGGCAAGTACTCTTTCAACTGGGCGGTGAGGACCAAGCTCAAACTCACTCCAA	9214
	QY	1628	TTCCGGCTGGCTCCCGGTGAGCTTGTCCGGCTGGTTCTGTTGTGGCTACAGCGGGGAG	1687
	DB	9215	TCCAGCGCGGTCCCGGTGGACTTGTCCGGCTGGTTCTGTTGTGTTACAGCGGGGAG	9274
	QY	1688	ACATATATCAGAGCTGTCTGTGGCCCGACCCCGC	1722
	DB	9275	ACATATATCAGAGCTGTCTGTGGCCCGACCCCGC	9309

```

RESULT 7
ADD93729
ID ID ADD93729 standard; DNA; 7979 BP.
XX XX
XX AC ADD93729;
XX XX
XX XX 29-JAN-2004 (first entry)
XX DT
XX DT Hepatitis C virus strain J4 (B/RI) replicon.
DE XX
XX XX HCV; vaccine; virucide; ss.
XX KW
XX XX Hepatitis C virus.
XX OS
XX XX WO2003085084-A2.
PN XX
XX XX 16-OCT-2003.
XX PD
XX PP
XX PP 03-APR-2003; 2003WO-US010177.
XX XX
XX PP 03-APR-2002; 2002US-0369685P.
XX XX
XX PA (SMIX ) SMITHKLINE BEECHAM CORP.

```

XX Gates A, Gu B, Sarisky RT;
 XI WPI; 2003-804301/75.
 XX
 XX New hepatitis C virus (HCV) sub-genomic replicon, useful for facilitating
 PT screening or testing of anti-HCV drugs, comprises a nucleic acid
 PT construct encoding chimeric HCV non-structural proteins, and an NS5B
 PT polymerase gene.
 XX
 XX Disclosure; Page 61-65; 159pp; English.
 PS
 XX
 CC The present sequence comprises a replicating hepatitis C virus (HCV) J4
 CC (B/781) replicon. The invention provides sub-genomic replicons of HCV
 CC comprising a nucleic acid construct encoding chimeric HCV nonstructural
 CC protein and an NS5B polymerase gene. A preferred replicon comprises an
 CC NS3 nucleotide sequence ADD93721 that encodes the first 75 contiguous N-
 CC terminal amino acids of the NS3 of genotype 1b, of a BB7 strain. A
 CC chimeric replicon may comprise an NS3 sequence from any of the 6 major
 CC HCV genotypes and subtypes but has its first 225 nucleotides of the
 CC coding sequence replaced by the BB7 strain NS3 sequence, especially where
 CC the replicon is from HCV genotype 1a (H77 strain) or genotype 1b (J4
 CC strain). Stable cell lines expressing and replicating functional
 CC replicons containing sequences from HCV genotype 1a (strain H77) or
 CC genotype 1b (strain J4) within the prototype 1b replicon backbone from
 CC HCV strain BB7 are provided. These can be used to screen for compounds
 CC that modulate viral replication. The sub-genomic HCV replicon systems of
 CC the invention may provide the foundation for generating HCV replicons of
 CC all 6 major genotypes and subtypes to facilitate screening, testing and
 CC evaluating anti-infective agents for HCV disease(s).
 XX
 XX Sequence 7979 BP; 1663 A; 2361 C; 2223 G; 1732 T; 0 U; 0 Other;

Query Match	88.9%	Score 1547	DB 10	Length 7979
Best Local Similarity	93.9%	Pred. No. 0		
Matches 1610	Conservative 0	Mismatches 105	Indels 0	Gaps 0
Qy	8	GCATGTCAAATGTCCTATACATGGACAGGGCGCCTGATCACACCTCGCGCTCGGAGGAAA	67	
Db	5978	GCTGCTCAATGTCCTATACGTGGACAGGGCGCCTGATCAGCCATCGCTCGGAGGAAA	603	
Qy	68	GCAAGCTGCCCATCAACGGCTGAGCAACTCTTGTGCTGGTCACCATAACTGGTCTATT	127	
Db	6038	GTAAGCTGCCCATCAACCCGCTTGAGCAACTCTTGTGCTGGTCACCAACAATGGTCTACG	609	
Qy	128	CCACAACATCCGCAAGTGAAGCTCGCGCAGAGAGAGGTTCACCTTTGACAGACTGCAG	187	
Db	6098	CCACAACATCCCGAGCGCAAGCTCGCGCAGAGAGAGGTTCACCTTTGACAGATTGCAAG	615	
Qy	188	TCCTGGACGATCAATACCGGGAGCGTCTCAAGGAGATGAAGGGCAAGCGCTGCACAGTGA	247	
Db	6158	TCCTGGATGATCAATACCGGACGTACTCAAGGAGATGAAGGGCAAGCGCTGCACAGTTA	621	
Qy	248	AGGCTAACTGCTATCTGTAGAAAGACATGCAAGCTGACGCCGCCCATTTGGGCCAAAT	307	
Db	6218	AGGCTAAGCTTCTATCTATAGAGGAGGCTGCAAGCTGACGCCGCCACATTTGGGCCAAAT	627	
Qy	308	CCAAATTTGGCTATGGGCAAGGACGTCGGAGCCTATCCAGCAGGGCGGTTAAACCACA	367	
Db	6278	CCAAATTTGGCTATGGGCAAGGACGTCGGAAACCTATCCAGCAGGGCGGTTAAACCACA	633	
Qy	368	TCCGCTCCGTGTGGGAAGGACTTGTCTGGAGACACTGACACACCAATTCAGACCACTCA	427	
Db	6338	TCCGCTCCGTGTGGGAGGACTTGTCTGGAAGACACTGAAACACCAATTTGACACCACTCA	639	
Qy	428	TGGCAAAAATGAGGTTTTCTGGGTCCAACGAGAGAAAGAGCGCGCAAAACCAAGCTCGCC	487	
Db	6398	TGGCAAAAATGAGGTTTTCTGGGTCCAACGAGAGAAAGAGCGCGCAAGCCAGCTCGCC	645	
Qy	488	TCATCGTATTCACAGACCTGGGAGTTTCGTGTATGCCAGAAAGATGGCCCTTTACGACGTGG	547	
Db	6458	TTATCGTATTCACAGACCTGGGAGTTTCGTGTATGCCAGAAAGATGGCCCTTTACGACGTGG	651	

QY 548 TTTCACACTCTTCTCAGGCCCGTGATGGGCTCCTCATACGGATTCCAATACCTCTCTTAAGC 607
Db 6518 TCTCCACCCTTCTCAGGCCCGTGATGGGCTCCTCATACGGATTCCAATACCTCTCTTAAGC 6577
QY 608 AGCGGGTCAGTTCTCGTGAATACCTGGAAGCAAAGAAATGCCCTATGGGCTTCTCAT 667
Db 6578 AGCGGGTCAGTTCTCGTGAATACCTGGAATCAAGAAATGCCCTATGGGCTTCTCAT 6637
QY 668 ATGACACCCGCTGTTTCTCACTCAACGGTCACTGAGATGATCGGTTCAGAGTCAA 727
Db 6638 ATGACACCCGCTGTTTCTCACTCAACGGTCACTGAGATGATCGGTTCAGAGTCAA 6697
QY 728 TTTACCAATGTTGACACTTGGCCCCCGAAGCTAGACAGGCCATTAAGGTCGCTCACAGAGC 787
Db 6698 TTTACCAATGTTGACACTTGGCCCCCGAAGCTAGACAGGCCATTAAGGTCGCTCACAGAGC 6757
QY 788 GGCTCTATGTGGGGGTCCCATGACTAACTCCAAAGGCGAGAAGCTGGGGCTATCGCCGGT 847
Db 6758 GGCTTTACATGGGGGTCCCTGACTAACTCAAAGGCGAGACTGGGGTTATCGCCGGT 6817
QY 848 GCGGCGAGCGCGGTGCTGACGACTAGCTGCGGTAAATACCCCTCACATGCTACTTGAAGG 907
Db 6818 GCGGCGAAGTGGGTGCTGACGACTAGCTGCGGTAAATACCCCTCACATGCTACTTGAAGG 6877
QY 908 CGCTGACGCTGTCGAGTGCAGCTCCAGCTCCAGGACTGCAGATGCTCGTGAATGGAGAGC 967
Db 6878 CCAGCTGCAGCTGTCGAGTGCAGCTCCAGGACTGCAGATGCTCGTGAATGGAGAGC 6937
QY 968 ACCTTGTGCTTATCTGTGAAGCGGGGAACCCAAAGAGAGCGCGCAAGCTTACAGTCT 1027
Db 6938 ACCTTGTGCTTATCTGTGAAGCGGGGAACCCAAAGAGAGTGGCGGCCCTACAGGCT 6997
QY 1028 TCACGAGAGCTATGACTAGGTACTCTGCCCCCTCGGGACCCGCCCAACCGGAATACG 1087
Db 6998 TCACGAGAGCTATGACTAGGTATTCGCGCCCCCGGGGATCCGCCCCCAACAGAAATACG 7057
QY 1088 ACTTGAGAGTGAATAACATCGTGTCTCCAAATGTGTGCTGGTGGCAACAGATGATCTGGCA 1147
Db 7058 ACTTGAGAGTGAATAACATCATGTTCTCCAAATGTGTGCTGGTGGCAACAGATGATCTGGCA 7117
QY 1148 AAGGGTGTACTACCTACCGCTGACCCCGTCCCGTGGGGGCTGGTGGGAGA 1207
Db 7118 AAGGGTGTACTACCTACCGCTGACCCCGTCCCGTGGGGGCTGGTGGGAGA 7177
QY 1208 CAGCTAGGACACTCCAGTCACTCTGCTGAGCAACATCATATGATGTCGCCACTT 1267
Db 7178 CAGCTAGACACTCCCAATCACTCTGCTGAGCAATATCATATGATGTCGCCACTT 7237
QY 1268 TGTGGCAAGGATGATCTGATGACTCACTCTTCTCCATCTTTAGCCAGGAGCAAC 1327
Db 7238 TATGGGCAAGGATGATCTGATGACTCACTTTTCTCCATCTTTCTAGCTCAAGAGCAAC 7297
QY 1328 TTGAAAAAGCCCTGGATGTCAGATCTACGGGCTTGTACTCCATTGAGCCACTTGACC 1387
Db 7298 TTGAAAAAGCCCTGGATGTCAGATCTACGGGCTTGTACTCCATTGAGCCACTTGACC 7357
QY 1388 TACTCAGATCATTTGAACGACTCCATGCTTACCGGCTTACCGCAATTTTCACTCCATAGTCTCTC 1447
Db 7358 TACTCAGATCATTTGAACGACTCCATGCTTACCGGCTTACCGCAATTTTCACTCCATAGTCTCTC 7417
QY 1448 CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAAGCTTGGGTACCAACCTTGCAGAG 1507
Db 7418 CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAACCTTGGGTACCAACCTTGCAGAG 7477
QY 1508 TCTGGAGACATCGGGCCAGAAATGTCGCGCTAAAGTTACTGTCCCGAGGGGGAGGGCGG 1567
Db 7478 CTTGGAGACATCGGGCCAGAAATGTCGCGCTAAAGTTACTGTCCCGAGGGGGAGGGCGG 7537
QY 1568 CCATTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAACTCACTCCAA 1627
Db 7538 CCATTTGTGGCAAGTACTCTTAACTGGGCGAGTAAGGACCAAGCTTAACTCACTCCAA 7597
QY 1628 TTCGGGCTGGCTCCCGGCTGAGCTTGTGCGGCTGGTTGCTGGCTACAGCGGGGGAG 1687

Db 7598 TCCCGCCCGCTCCCGAGCTTGCTGCTGGTTCTGCTGTACAGCGGGGAG 7657
QY 1688 ACATATATCACAGCTGTCTCGTGCCCGACCCCGC 1722
Db 7658 ACATATATCACAGCTGTCTCGTGCCCGACCCCGC 7692
RESULT 8
ADD93730
ID ADD93730 standard; DNA; 7979 BP.
XX
AC ADD93730;
XX
XX 29-JAN-2004 (first entry)
XX Hepatitis C virus strain J4 (J4B/R1(C)) replicon.
XX HCV; vaccine; virucide; ss.
XX Hepatitis C virus.
XX W02003085084-A2.
XX 16-OCT-2003.
XX 03-APR-2003; 2003WO-US010177.
XX 03-APR-2002; 2002US-0369685P.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Gates A, Gu B, Sarisky RT;
XX WPI; 2003-804301/75.
XX
XX New hepatitis C virus (HCV) sub-genomic replicon, useful for facilitating screening or testing of anti-HCV drugs, comprises a nucleic acid construct encoding chimeric HCV non-structural proteins, and an NS5B polymerase gene.
XX
XX Disclosure; Page 65-70; 159pp; English.
XX
XX The present sequence comprises a replicating hepatitis C virus (HCV) J4 (J4B/R1(C)) replicon. The invention provides sub-genomic replicons of HCV comprising a nucleic acid construct encoding chimeric HCV nonstructural protein and an NS5B polymerase gene. A preferred replicon comprises an NS3 nucleotide sequence ADD93721 that encodes the first 75 contiguous N-terminal amino acids of the NS3 of genotype 1b of a BB7 strain. A chimeric replicon may comprise an NS3 sequence from any of the 6 major HCV genotypes and subtypes but has its first 225 nucleotides of the coding sequence replaced by the BB7 strain NS3 sequence, especially where the replicon is from HCV genotype 1a (H77 strain) or genotype 1b (J4 strain). Stable cell lines expressing and replicating functional replicons containing sequences from HCV genotype 1a (strain H77) or genotype 1b (strain J4) within the prototype 1b replicon backbone from HCV strain BB7 are provided. These can be used to screen for compounds that modulate viral replication. The sub-genomic HCV replicon systems of the invention may provide the foundation for generating HCV replicons of all 6 major genotypes and subtypes to facilitate screening, testing and evaluating anti-infective agents for HCV disease(s).
XX
XX Sequence 7979 BP; 1662 A; 2361 C; 2223 G; 1733 T; 0 U; 0 Other;
Query Match 88.9%; Score 1547; DB 10; Length 7979;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1610; Conservative 0; Mismatches 105; Indels 0; Gaps 0

QY 8 GCATGCTCAATGTCCTATACATGACAGCGCCCTGATCACACCGTCCGCGAGGAAA 67
Db 5978 GCTGCTCAATGTCCTATACGTTGGACAGCGCCCTGATCACGCCATGCTCGGAGGAAA 603
QY 68 GCAAGCTGCCCATCAACGCGCTGAGCAACTCTCTTCTGCTGCCATCAACCTGCTATT 127

Db	6038	GTAAAGCTGCCCATCAACCCGTTGAGCAACTTTTCTCTCGGTCAACCAACATGGTCTACG	6097
QY	128	CCACAACATCCCGCAGTGCAGCCCTCGGCAGAAAGGTCACTTTTGACAGATGCAAG	187
Db	6098	CCACAACATCCCGCAGCGCAAGCCCTCGGCAGAAAGGTCACTTTTGACAGATGCAAG	6157
QY	188	TCTTGACCATCATTAACGGGACGTGCTCAAGGAGATGAAGGCGGAAGCGCTCCACAGTGA	247
Db	6158	TCTGTGATGATCATTTACGGGACGTACTCAAGGAGATGAAGCGCAGAGCGCTCCACAGTTA	6217
QY	248	AGGCTAACTGCTATCTGTAGAAAGCATGCAAGCTGAACGCCCCCGCATTTCCGCCAAAT	307
Db	6218	AGGCTAAGCTTCTATCTATAGAGGAGCGCTCAAGCTGACGCCCCACATTTCCGCCAAAT	6277
QY	308	CCAAATTTGGCTATCGGGCAAGGACGTCGAGGCGCTATCAGCAGAGCGCGTTAAACACA	367
Db	6278	CCAAATTTGGCTATCGGGCAAGGACGTCGAGGCGCTATCAGCAGAGCGCGCTTAAACACA	6337
QY	368	TCCGCTCCGCTGTGGAAGGACTTGTCTGGAGGACACTGACACACCAATTCAGACCACCATCA	427
Db	6338	TCCGCTCCGCTGTGGAAGGACTTGTCTGGAGGACACTGAAACACCAATTTGACACCAACATCA	6397
QY	428	TGGCAAAAATGAGGTTTTCTGCGTCCAAACAGAGAAAGGAGCGCGAAACACAGCTCGCC	487
Db	6398	TGGCAAAAAGTGAGGTTTTCTGCGTCCAAACAGAGAAAGGAGCGCGCAAGCTCGCC	6457
QY	488	TCATCGTATTCACAGACTGGGAGTTCTGTATCGGAGAGATGGCCCTTTACGACGTGG	547
Db	6458	TTATCGTATTCACAGACTGGGAGTTCTGTATCGGAGAGATGGCCCTTTACGACGTGG	6517
QY	548	TTTTCACACTTCTCTCAGGCCGTGATGGGCTCCTCATACGGATTCCAATCTCTCTAAGC	607
Db	6518	TCCTCACCTTCTCTCAGGCCGTGATGGGCTCCTCATACGGATTCCAATCTCTCTAAGC	6577
QY	608	AGCGGGTCAGTTCTCTGGTGAATACCTTGGAGCAAGAAATGCCCTATCGGCTTCTCAT	667
Db	6578	AGCGGGTCAGTTCTCTGGTGAATACCTTGGAGCAAGAAATGCCCTATCGGCTTCTCAT	6637
QY	668	ATGACACCCCGCTGTTTGTACTCAACGGTCACTGAGAAATGATCCGTTGTGAGGAGTCAA	727
Db	6638	ATGACACCCCGCTGTTTGTACTCAACGGTCACTGAGAAATGATCCGTTGTGAGGAGTCAA	6697
QY	728	TTTACCATAATGTTGTACTTGGCCCCCGAAGCTAGACAGGCCATAAGGTGCTCCACAGAGC	787
Db	6698	TTTACCATAATGTTGTACTTGGCCCCCGAAGCCAGACAGGCCATAGGTGCTCCACAGAGC	6757
QY	788	GGCTCTATGTCCGGGGTCCCATGACTAACTCCAAGGGCAGAACTCGCGGTATTCGCCGT	847
Db	6758	GGCTTTACATCCGGGGTCCCTGACTAACTCAAAAGGGCAGAACTCGCGGTATTCGCCGT	6817
QY	848	CCCGGCGAGCGGCTGCTGACGACTAGCTGCGGTAAATACCTCACATGCTACTTTGAAGG	907
Db	6818	CCCGGCGAAGTGGGCTGCTGACGACTAGCTGCGGTAAATACCTCACATGCTACTTTGAAGG	6877
QY	908	CCGCTGCAGCTGTGAGACTGCCAAGCTCCAGGATGACGATGCTGTGAATGAGAGCG	967
Db	6878	CCACTGCAGCTGTGAGACTGCCAAGCTCCAGGATGACGATGCTGTGAATGAGAGCG	6937
QY	968	ACCTTGTCTTATCTGTGAAGCGCGGGAACCCAGAGGACGCGGCAAGCCTACAGTCT	1027
Db	6938	ACCTTGTCTTATCTGTGAAGCGCGGGAACCCAGAGGATGCGGCGGCTTACAGGCT	6997
QY	1028	TCACGGAGGCTATGACTAGGTACTCTGCGCCCGCTCGGGGACCGCGCCCAACCGGAATACG	1087
Db	6998	TCACGGAGGCTATGACTAGGTACTCTGCGCCCGCTCGGGGATCGCGCCCAACCGAATACG	7057
QY	1088	ACTTGGAGCTGATAAATATCGTGTCTCTCAATGTGCGGTGCGACATGCTATCGGCA	1147
Db	7058	ACCTGGAGCTGATAAATATCGTGTCTCTCAATGTGTCAGTTCGCGCAGATGCTATCGGCA	7117
QY	1148	AAAGGGTGTACTACCTCACCGCGTGAACCCACCGTCCCTTTCGCGGGGTGTCGTGGGAGA	1207

XX The present sequence comprises a replicating hepatitis C virus (HCV) J4
CC replicon. The invention provides sub-genomic replicons of HCV comprising
CC a nucleic acid construct encoding chimeric HCV nonstructural protein and
CC an NS5B polymerase gene. A preferred replicon comprises an NS3 nucleotide
CC sequence ADD91721 that encodes the first 75 contiguous N-terminal amino
CC acids of the NS3 of genotype 1b, of a B97 strain. A chimeric replicon may
CC comprise an NS3 sequence from any of the 6 major HCV genotypes and
CC subtypes but has its first 225 nucleotides of the coding sequence
CC replaced by the B97 strain NS3 sequence, especially where the replicon is
CC from HCV genotype 1a (H77 strain) or genotype 1b (J4 strain). Stable cell
CC lines expressing and replicating functional replicons containing
CC sequences from HCV genotype 1a (strain H77) or genotype 1b (strain J4)
CC within the prototype 1b replicon backbone from HCV strain B97 are
CC provided. These can be used to screen for compounds that modulate viral
CC replication. The sub-genomic HCV replicon systems of the invention may
CC provide the foundation for generating HCV replicons of all 6 major
CC genotypes and subtypes to facilitate screening, testing and evaluating
CC anti-infective agents for HCV disease(s).

XX
SQ Sequence 7979 BP; 1673 A; 2351 C; 2216 G; 1739 T; 0 U; 0 Other;

Query Match 88.9%; Score 1547; DB 10; Length 7979;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1610; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY	8	GCATGTCAATGCTCTATACATGACAGCGCGCCTGATCACACCGTGCCTCGGAGGAAA	67
DB	5978	GCTGTCAATGCTCTATACGTGGACGGCGCCTGATCACGCCATGCGCTCGGAGGAAA	6037
QY	68	GCAAGCTGCCATCAACCGCTGAGCAACTCTTCTGCTGGCTCACCATTAACCTGTCTATT	127
DB	6038	GTAAGCTGCCATCAACCGTTGAGCAACTCTTCTGCTGCTCACCAACATGCTCTACG	6097
QY	128	CCACAACATCCCGAGTGGACGCTTGGCGAGCAAGAGTCACTTTGACAGACTGCGAAG	187
DB	6098	CCACAACATCCCGAGCGCAAGCCTCCGCGAGCAAGAGTCACTTTGACAGATTGCGAAG	6157
QY	188	TCTGGACCATCATTTACCGGGACGCTGCTCAAGAGATGAAGCGAAGCGTCCACAGTGA	247
DB	6158	TCTGTGATGATCATTTACCGGGACGCTGCTCAAGAGATGAAGCGAAGCGTCCACAGTGA	6217
QY	248	AGGCTAAATGCTATCTGTAGAAGAGCATGCAAGCTGACGCGCCCGCATTCGGCCAAAT	307
DB	6218	AGGCTAAAGCTTCTATCTATAGAGGAGCGCTGCAAGCTGACGCGCCCGCATTCGGCCAAAT	6277
QY	308	CCAAATTTGGCTATGGGGCAAGAGAGTCCGGAGCCTATCCAGAGGGCGGTTTACCACA	367
DB	6278	CCAAATTTGGCTATGGGGCAAGAGAGTCCGGAAACCTATCCAGAGGGCGGTTTACCACA	6337
QY	368	TCCGCTCCGTTGGAGGACTTGTCTGGAGGACACTGACACACCAATTCAGACCACTCA	427
DB	6338	TCCGCTCCGTTGGAGGACTTGTCTGGAGGACACTGAAACACCAATTTGACACCACTCA	6397
QY	428	TGCGAAAAATAGGTTTCTGGCTCCAAACAGAGAAAGAGCGCGCAACACAGCTCGCC	487
DB	6398	TGCGAAAAATAGGTTTCTGGCTCCAAACAGAGAAAGAGCGCGCAACACAGCTCGCC	6457
QY	488	TCACTCGTATTTCCAGACCTGGGAGTTTCTGTATGCGAGAAGATGGCCCTTTTACGCTGG	547
DB	6458	TATTCGTATTTCCAGACCTGGGAGTTTCTGTATGCGAGAAGATGGCCCTTTTACGCTGG	6517
QY	548	TTTCCCACTTCTTCCAGGCGTGATGGGCTCTCTATACGGATCCCAATCTCTCTTAAGC	607
DB	6518	TCTCCACCTTCTCTCAGGCGTGATGGGCTCTCTATACGGATTTCAATATCTCCCCCAAGC	6577
QY	608	AGCGGGTCGATTTCTGTTGAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT	667
DB	6578	AGCGGGTCGATTTCTGTTGAATACCTGGAAATCAAGAAATGCCCTATGGGCTTCTCAT	6637
QY	668	ATGACACCCGCTGTTTGTAGCTCAACGCTCACTGAGAAATGACATCCCGTTGTGAGGAGTCAA	727
DB	6638	ATGACACCCGCTGTTTGTAGCTCAACGCTCACTGAGAGTGACATTCGTTGTGAGGAGTCAA	6697

QY	728	TTTACCAATGTTGTGACTTTGGCCCGGAAAGCTAGACAGGCCATAAGGTCGCTCACAGAGC	787
DB	6698	TTTACCAATGTTGTGACTTTGGCCCGGAGGCCAGACAGGCCATAAGGTCGCTCACAGAGC	6757
QY	788	GGCTCTATGTCGGGGGTCCATGACTAACTCCAAAGGGCAGAACTCGCGCTATTCGCGGCT	847
DB	6758	GGCTTTTACATCGGGGGTCCCTGACTTAACTCAAAAAGGGCAGAACTCGCGTTATTCGCGGCT	6817
QY	848	GCCGCGGAGCGCGCTGCTGACGACTAGCTGGGTAAATACCTCCACATGCTACTTCAAGG	907
DB	6818	GCCGCGCAAGTGGCGTCTGACGACTAGCTGGGGTAAATACCTCCACATGCTACTTCAAGG	6877
QY	908	CCGCTCGAGCTGCTCGAGCTGCCAAGCTCCAGACTGCAAGCTGCTCGTGAATGGAGAGC	967
DB	6878	CCACTGCGAGCTGCTCGAGCTGCCAAGCTCCAGACTGCAAGCTGCTCGTGAAGCGAGAGC	6937
QY	968	ACCTTGTGCTTATCTGTGAAAGCGCGGAAACCCCAAGAGGACGCGCAAGCTTACGAGTCT	1027
DB	6938	ACCTTGTGCTTATCTGTGAAAGCGCGGAAACCCAGAGGATGCGCGGCGCTTACGAGCCT	6997
QY	1028	TCACGAGGCTATGACTAGGTACTCTTGCCCCCTCGGGNACCCGCCACCGGAATACG	1087
DB	6998	TCACGAGGCTATGACTAGGTATTTCCGCCCTCCCGGGGATCCGCCCAACCCAGAAATACG	7057
QY	1088	ACTTGGAGCTGATAACATCGTGTCTCCAAATGTGCTGCGCACACGATGCAATCTGGCA	1147
DB	7058	ACCTGGAGCTGATAACATCAATGTTCTTCCAAATGTGTGCTGCGCACGATGCAATCTGGCA	7117
QY	1148	AAAGGGTGTACTACTCCACCCGTGACCCCAACCCGCTCCCTTTGGCGGGCTGCGTGGGAGA	1207
DB	7118	AAAGGGTGTACTACTCCACCCGTGACCCCAACCCGCTCCCTTTGGCACGGCTGCGTGGGAGA	7177
QY	1208	CAGCTAGGCACACTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGCGGCCACTT	1267
DB	7178	CAGCTAGGCACACTCCAATCAACTCTTGGCTAGGCAATATCATCATGTATGCGGCCACCC	7237
QY	1268	TGTGGGCAAGATGATTTCTGATGACTCACTTCTTCTCATCTCTTAGCCACAGGAGCAAC	1327
DB	7238	TATGGGCAAGATGATTTCTGATGACTCACTTTTCTCATCTCTTAGCTCAAGAGCAAC	7297
QY	1328	TTGAAAAGCCCTCGATTTGTGAGTCTACGGGCTTGTACTTCTCATCTCTTAGCCACTTGACC	1387
DB	7298	TTGAAAAGCCCTCGATTTGTGAGTCTACGGGCTTGTACTTCTCATCTCTTAGCCACTTGACC	7357
QY	1388	TACCTCAGATCATTTGAACGACTCCATGCTCTTAGCGCATTTACACTCCACAGTTACTCTC	1447
DB	7358	TACCTCAGATCATTTGAACGACTCCATGCTCTTAGCGCATTTACACTCCACAGTTACTCTC	7417
QY	1448	CAGGTGAGATCAATAGGGTGGCTTCATGCTCTAGGAAGCTTTGGGGTACCACCTTTGGAA	1507
DB	7418	CAGGTGAGATCAATAGGGTGGCTTCATGCTCTAGGAAGCTTTGGGGTACCACCTTTGGAA	7477
QY	1508	TCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTTACTGTCTCCACGGGGGAGGGCGC	1567
DB	7478	CCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTTACTGTCTCCACGGGGGAGGGCGC	7537
QY	1568	CCATTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAA	1627
DB	7538	CCATTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAA	7597
QY	1628	TTCCGGCTCGCTCCCGGCTGGACTTGTCCGCGTGGTTGTGTTGCTGGGTACAGCGGGGAG	1687
DB	7598	TCCCGGCCGCTCCAGCTGGACTTGTCTGGCTGGTTGTCTGGCTGGTTGTACAGCGGGGAG	7657
QY	1688	ACATATATACAGGCTGTCTGTCGCCGACCCCGC	1722
DB	7658	ACATATATACAGGCTGTCTGTCGCCGACCCCGC	7692

RESULT 10
ADD93731
ID ADD93731 standard; DNA; 7979 BP.

XX ADD93731;
XX 29-JAN-2004 (first entry)
XX Hepatitis C virus strain J4 replicon.
XX HCV; vaccine; virucide; ss.
XX
XX Hepatitis C virus; genotype 1b.
XX WO2003085084-A2.
XX 16-OCT-2003.
XX 03-APR-2003; 2003WO-US010177.
XX 03-APR-2002; 2002US-0369685P.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Gates A, Gu B, Sarisky RT;
XX WPI; 2003-804301/75.
XX
XX New hepatitis C virus (HCV) sub-genomic replicon, useful for facilitating
XX screening or testing of anti-HCV drugs, comprises a nucleic acid
XX construct encoding chimeric HCV non-structural proteins, and an NS5B
XX polymerase gene.
XX
XX Disclosure; Page 70-75; 159pp; English.
XX
XX The present sequence comprises a replicating hepatitis C virus (HCV) J4
XX replicon. The invention provides sub-genomic replicons of HCV comprising
XX a nucleic acid construct encoding chimeric HCV nonstructural protein and
XX an NS5B polymerase gene. A preferred replicon comprises an NS3 nucleotide
XX sequence ADD93721 that encodes the first 75 contiguous N-terminal amino
XX acids of the NS3 of genotype 1b, of a BB7 strain. A chimeric replicon may
XX comprise an NS3 sequence from any of the 6 major HCV genotypes and
XX subtypes but has its first 225 nucleotides of the coding sequence
XX replaced by the BB7 strain NS3 sequence, especially where the replicon is
XX from HCV genotype 1a (H77 strain) or genotype 1b (J4 strain). Stable cell
XX lines expressing and replicating functional replicons containing
XX sequences from HCV genotype 1a (strain H77) or genotype 1b (strain J4)
XX within the prototype 1b replicon backbone from HCV strain BB7 are
XX provided. These can be used to screen for compounds that modulate viral
XX replication. The sub-genomic HCV replicon systems of the invention may
XX provide the foundation for generating HCV replicons of all 6 major
XX genotypes and subtypes to facilitate screening, testing and evaluating
XX anti-infective agents for HCV disease(s).
XX
XX Sequence 7979 BP; 1674 A; 2351 C; 2216 G; 1738 T; 0 U; 0 Other;
XX
XX Query Match 88.9%; Score 1547; DB 10; Length 7979;
XX Best Local Similarity 93.9%; Pred. No. 0;
XX Matches 1610; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
XX
XX 8 GCATGTCAATGCTCTATACATGACGCGCCCTGATCACACGTCGCGTCCGAGGAAA 67
XX
XX 5978 GCTGCTCAATGTCTATACGTGGACAGCGCCCTGATCGCCATCGCGTGGAGAAA 6037
XX
XX 68 GCAAGCTGCCATCAACGCGCTGAGCAACTCTTGTGCGTCAACATAACTGCTCTATT 127
XX
XX 6038 GTAGCTGCCATCAACCCGTTGAGCAACTTTTGTGCGTCAACACATGCTCTACG 6097
XX
XX 128 CCACACATCCGCGAGTGCAGCCTCGCGCAGAGAGAGGTCACTTTGACAGACTGCAAG 187
XX
XX 6098 CCACACATCCGCGAGCAGCCTCCGCGCAGAGAGAGGTCACTTTGACAGATTGCAAG 6157
XX
XX 188 TCCTGGACGATCATTTACCGGAGCGTCTCAAGGAGATGAAGCGAGCGGCTCCACAGTGA 247
XX
XX 6158 TCCTGGATGATCATTTACCGGAGCGTACTCAAGGAGATGAAGCGAGCGGCTCCACAGTTA 6217

QY 248 AGGCTAAACTGCTATCTGTAGAAGACATGCAAGCTGAOCGCCCGGCAATTTCGGCCAAAT 307
DB |||||
QY 6218 AGGCTAAGCTTCTATCTATAGAGGAGGCTTGAAGCTGACGCCCCACATATTCGGCCAAAT 6277
DB |||||
QY 308 CCAAAATTTGGCTATGGGCAAGGACGTCCCGAGGCTATCCAGCAGGCGCGTTAAACCACA 367
DB |||||
QY 6278 CCAAAATTTGGCTATGGGCAAGGACGTCCCGAGGCTATCCAGCAGGCGCGTTAAACCACA 6337
DB |||||
QY 368 TCCGCTCCGCTGGAAGGACCTTGTGAGGACACTGACACACCAATTCAGACCCACCATCA 427
DB |||||
QY 6338 TCCGCTCCGCTGGAAGGACCTTGTGAGGACACTGGAACACCAATTCAGACCCACCATCA 6397
DB |||||
QY 428 TGGCAAAAATGAGGTTTCTGCGTCCACAGAGAAAGAGCGCGCAAAATCAGTCTCGCC 487
DB |||||
QY 6398 TGGCAAAAATGAGGTTTCTGCGTCCACAGAGAAAGAGCGCGCAAAATCAGTCTCGCC 6457
DB |||||
QY 488 TCATCGTATTCCTCCAGACCTCGGAGTTCGTGTATCCGAGAAATGCGCCCTTTTACGACGTGG 547
DB |||||
QY 6458 TTATCGTATTCCTCCAGACCTCGGAGTTCGTGTATCCGAGAAATGCGCCCTTTTACGACGTGG 6517
DB |||||
QY 548 TTTCCACTCTTCTCAGCCCGTGTATGCGCTCTCATACGGAATTCAAATCTCTCTTAAGC 607
DB |||||
QY 6518 TCTCCACCTTCTCAGCCCGTGTATGCGCTCTCATACGGAATTCAAATCTCTCTTAAGC 6577
DB |||||
QY 608 AGCGGTCGAGTTCTCGTGAATACCTTGAAGCAAAAGAAATGCGCTTATGCGCTTCTCAT 667
DB |||||
QY 6578 AGCGGTCGAGTTCTCGTGAATACCTTGAAGCAAAAGAAATGCGCTTATGCGCTTCTCAT 6637
DB |||||
QY 668 ATGACACCCGCTGTTTGTGACTCAACGCTCACTGAGATGACATCCGTTGTGAGAGTCAA 727
DB |||||
QY 6638 ATGACACCCGCTGTTTGTGACTCAACGCTCACTGAGATGACATTCGTTGTGAGAGTCAA 6697
DB |||||
QY 728 TTTACCAATGTTGACTTGGCCCCCGAAGCTACAGAGGCAATAGCTCGCTCACAGAGC 787
DB |||||
QY 6698 TTTACCAATGTTGACTTGGCCCCCGAAGCTACAGAGGCAATAGCTCGCTCACAGAGC 6757
DB |||||
QY 788 GGCTCTATGTGCGGGGTCCCATGACTAACTCAAGGCGCAGAACTGCGGCTATCGCCGCT 847
DB |||||
QY 6758 GGCTCTATGACTGCGGGGTCCCATGACTAACTCAAGGCGCAGAACTGCGGCTATCGCCGCT 6817
DB |||||
QY 848 GCGCGGAGCGGGGTGCTGACGACTAGCTGCGGTAAATACCTCAATGCTTACTTTGAAGG 907
DB |||||
QY 6818 GCGCGCAAGTGGCGGTGCTGACGACTAGCTGCGGTAAATACCTCAATGCTTACTTTGAAGG 6877
DB |||||
QY 908 CCGCTCAGAGCTGCTCGAGCTGCCAAGCTCCAGAGCTGACAGATGCTCGTGAATGAGAGC 967
DB |||||
QY 6878 CCGCTCAGAGCTGCTCGAGCTGCCAAGCTCCAGAGCTGACAGATGCTCGTGAATGAGAGC 6937
DB |||||
QY 968 ACCTTGTGTTATCTGTGAAAGCGCGGAAACCAAGAGGACGCGCAAGCTTACGAGTCT 1027
DB |||||
QY 6938 ACCTTGTGTTATCTGTGAAAGCGCGGAAACCAAGAGGATGCGGCGGCTTACGAGCCT 6997
DB |||||
QY 1028 TCAGGAGGCTATGACTAGTACTCTGCCCCCTTGGGAGCCGCCCAACCGGAATACG 1087
DB |||||
QY 6998 TCAGGAGGCTATGACTAGTACTCTGCCCCCTTGGGAGCCGCCCAACCGGAATACG 7057
DB |||||
QY 1088 ACTTGAGCTGATACATCGTGTCTCAATGTGTGCGTGCACACGATGCACTCTGGCA 1147
DB |||||
QY 7058 ACTTGAGCTGATACATCGTGTCTCAATGTGTGCGTGCACACGATGCACTCTGGCA 7117
DB |||||
QY 1148 AAAGGGTGTACTACCTCACCGGTGACCCCGCTCCCTTGGCGGGGTGCGTGGGAGA 1207
DB |||||
QY 7118 AAAGGGTGTACTACCTCACCGGTGACCCCGCTCCCTTGGCGGGGTGCGTGGGAGA 7177
DB |||||
QY 1208 CAGCTAGGCACACTCCAGTCAACTCTGCTGCTAGGCAACATCATCATGTATGCGCCACITT 1267
DB |||||
QY 7178 CAGCTAGGCACACTCCAGTCAACTCTGCTGCTAGGCAACATCATCATGTATGCGCCACITT 7237
DB |||||
QY 1268 TGTGGCAAGGATGATTTCTGATGACTCACTTTCTTCTCACTCTTCTAGCCCGAGGAGCAAC 1327
DB |||||
QY 7238 TATGGCAAGGATGATTTCTGATGACTCACTTTTCTCCATCTTCTTAGCTCAAGAGCAAC 7297
DB |||||
QY 1328 TTGAAAAGCCCTGGATTGTGATGACTACGCGGGCTTGTACTCCATTTGAGCCACTTGACC 1387
DB |||||

Db 7298 TTGAAAAGCCCTGGATTGTTCAGATCTACGGGGCTTGCTACTCCATTGAGCCACTTGACC 7357
Qy 1388 TACCTTCAGATCATTTGAAGCACTCCATGGTCTTAGCGGATTTTCACTCCATAGTTACTCTC 1447
Db 7358 TACCTTCAGATCATTTGAAGCACTCCATGGTCTTAGCGGATTTTCACTCCATAGTTACTCTC 7417
Qy 1448 CAGGTGAGATCAATAGGTGGCTTTCATGCTCCAGAAAGCTTGGGGTACCACTTTGGGAG 1507
Db 7418 CAGGTGAGATCAATAGGTGGCTTTCATGCTCCAGAAACTTTGGGGTACCACTTTGGGAA 7477
Qy 1508 TCTGAGACATCGGCCAGAGTGTCCCGCTTAAGTTACTGTCCAGGGGGGGAGGCGG 1567
Db 7478 CTTGAGACATCGGCCAGAGTGTCCCGCTTAAGTTACTGTCCAGGGGGGGAGGCGG 7537
Qy 1568 CCATTGTGGCAAGTACTCTTCACTGGGCGTAAGGACCAAGCTTAACTCACTCAA 1627
Db 7538 CCATTGTGGCAAGTACTCTTCACTGGGCGTAAGGACCAAGCTTAACTCACTCAA 7597
Qy 1628 TTCCGGCTCGTCCCGGTGGACTTGTCCGGCTGGTTCGTTGCTGGCTACAGCGGGGAG 1687
Db 7598 TCCCGGCGCGTCCCGGTGGACTTGTCCGGCTGGTTCGTTGCTGGCTACAGCGGGGAG 7657
Qy 1688 ACATATATCACAGCCTGTCTCGTCCCGGACCCCGC 1722
Db 7658 ACATATATCACAGCCTGTCTCGTCCCGGACCCCGC 7692

RESULT 11
AA24843
ID AAX24843 standard; DNA; 9595 BP.
XX AAX24843;
XX DT
XX 21-JUN-1999 (first entry)
XX Infectious hepatitis C virus genotype 1b strain HC-J4 genome.
DE
XX
DE HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;
KW assay; antiviral; virucide; ss.
XX
XX Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 342..9374
FT /*tag= a
XX
XX WO9904008-A2.
PN
XX
XX 28-JAN-1999.
PD
XX
XX 16-JUL-1998; 98WO-US014688.
PF
XX
XX 18-JUL-1997; 97US-0053062P.
PR 27-JAN-1998; 98US-00014416.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX WPI; 1999-132252/11.
DR P-PSDB; AAW98022.
XX
XX New isolated hepatitis C virus nucleic acids - used to develop products
PT for the diagnosis, prevention and treatment of HCV infections and for
PT developing screening assays.
XX
XX Claim 3; Fig 14A-F; 126pp; English.
XX
XX The present sequence comprises the nucleic acid sequence of the genome of
CC infectious hepatitis C virus (HCV) genotype 1b strain HC-J4 (ATCC 209596)
CC that is capable of expressing this virus when transfected into cells. HC-
CC J4 was obtained from acute phase plasma of a chimpanzee experimentally

CC infected with serum containing HC-J4/91. The claimed infectious nucleic
CC acid sequence can be used to produce chimeric genomes (see AAX248433)
CC consisting of the open reading frames of infectious nucleic acid
CC sequences of other genotypes (including genotypes 1-6) and subtypes (such
CC as 1b, 2a, 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also
CC relates to the introduction of mutations or deletions into infectious
CC nucleic acid sequences in order to produce an attenuated HCV virus
CC suitable for vaccine development. Infectious nucleic acid sequences can
CC also be used to produce attenuated virus of a host cell with the
CC of the viruses produced by transfection of a host cell with the
CC infectious nucleic acid sequence. Vaccines comprising one or more
CC polypeptides made from the infectious nucleic acid sequence are used to
CC immunise mammals, especially humans, against hepatitis C. The nucleic
CC acid sequences can also be used to induce protective immunity against the
CC virus. The nucleic acid sequences or their encoded proteases (e.g. NS3
CC protease) can additionally be used to develop screening assays to
CC identify antiviral agents for HCV
XX
SQ Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;
Query Match 88.9%; Score 1547; DB 2; Length 9595;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1610; Conservative 0; Mismatches 105; Indels 0; Gaps 0,
Qy 8 GCATGTCAATGTCTTATACATGACAGCGCCCTGATCACACCGTCGCGTCGCGAGGAAA 67
Db 7594 GCTGCTCAATGTCTTATACATGACAGCGCCCTGATCACACCGTCGCGTCGCGAGGAAA 7653
Qy 68 GCTAGCTGCCCATCAACGCGCTGAGCAACTCTTGTGCTGCCTACCAATAA.CCTGCTCTATT 127
Db 7654 GTAAGCTGCCCATCAACCGCTTGAACACTCTTGTGCTGCCTACCAATAA.CATGCTCTACG 7713
Qy 128 CCACAAACATCCCGCAGTGCAGGCTCGCGCAGAGAAAGTTCACCTTTGACAGACTTGAAG 187
Db 7714 CCACAAACATCCCGCAGGCGCAAGCTCTCGCGCAGAGAAAGTTCACCTTTGACAGACTTGAAG 7773
Qy 188 TCCTGGACGATCATATACCGGACGTCTCAAGGAGATGAAGGCGGAGCGCTCCACAGTGA 247
Db 7774 TCCTGGATGATCATATACCGGACGTCTCAAGGAGATGAAGGCGGAGCGCTCCACAGTGA 7833
Qy 248 AGGCTAAACTGCTATCTGTAGAGAAAGCATGCAAGCTGACGCCCGCCCATTTCCGCCAAAT 307
Db 7834 AGGCTAAAGTCTTCTATCTATAGAGAGGCGCTGCAAGCTGACGCCCGCCCATTTCCGCCAAAT 7893
Qy 308 CCAAAATTTGGCTATGGGCGCAAAAGACGTCGCGAGCGCTATCCAGCAGGCGCGTTAAACACA 367
Db 7894 CCAAAATTTGGCTATGGGCGCAAAAGACGTCGCGAGCGCTATCCAGCAGGCGCGTTAAACACA 7955
Qy 368 TCCGCTCCGCTGGAAGGACTTGTGAGGAGACATGACACACCAATTCAGACACCATCA 427
Db 7954 TCCGCTCCGCTGGAAGGACTTGTGGAAGACATGAAACACCAATTCAGACACCATCA 8012
Qy 428 TGGCAAAAATGAGGTTTCTGCGTCCAAACAGAGAAAGGAGGCGGCAACAGCTGCC 487
Db 8014 TGGCAAAAATGAGGTTTCTGCGTCCAAACAGAGAAAGGAGGCGGCAACAGCTGCC 8077
Qy 488 TCATCGTATTCACAGCCTGGAGTTTCTGCTATCGAGAGATGCGCCCTTTACGACGTGG 547
Db 8074 TTATCGTATTCACAGCCTGGAGTTTCTGCTATCGAGAGATGCGCCCTTTACGACGTGG 813
Qy 548 TTTCACCTCTTCTCCAGCGCTGATGGGCTCTCATACCGGATTTCAATACTCTCTTAAGC 607
Db 8134 TCTCCACCTCTTCTCCAGCGCTGATGGGCTCTCATACCGGATTTCAATACTCTCTTAAGC 619
Qy 608 AGCGGGTCCAGTTCTCTGGTGAATACCTGGAAAGCAAAAGAAATGCGCTATGGGCTTCTCAT 667
Db 8194 AGCGGGTCCAGTTCTCTGGTGAATACCTGGAAATCAAGAAATGCGCTATGGGCTTCTCAT 825
Qy 669 ATGACACCGCGTGTGTTGACTCAACGGTCACTGAGATGACATCGTCTGAGGAGTCAA 727
Db 8254 ATGACACCGCGTGTGTTGACTCAACGGTCACTGAGAGTGAACATCGTGTGAGGAGTCAA 831
Qy 728 TTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCCCATAGGTCGCTCACAGAGC 787

[illegible]

RESULT 12
AAF23492
ID AAF2:
XX
AC AAF2:

XX	21-MAR-2001 (first entry)
DT	
XX	
DE	Infectious Hepatitis C virus 1b genotype.
XX	
KW	GBV-B; hepatitis C virus; HCV; vaccine; ds.
XX	
OS	Hepatitis C virus.
XX	
PN	WO200075337-A1.
XX	
PD	14-DEC-2000.
XX	
PF	02-JUN-2000; 2000WO-US015293.
XX	
PR	04-JUN-1999; 99US-0137694P.
XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Bukh J, Yanagi M, Emerson SU, Purcell RH;
XX	
DR	WPI; 2001-091214/10.
XX	
PT	New infectious nucleic acids of the GB virus-B clone, useful for
PT	indirectly studying the molecular properties of hepatitis C virus (HCV)
PT	and in developing vaccines and therapeutics for HCV.
XX	
PS	Disclosure; Fig 7; 96pp; English.
XX	
CC	The present invention relates to GB virus-B. The nucleic acid molecules
CC	of the invention are useful for indirectly studying the molecular
CC	properties of hepatitis C virus (HCV). The infectious nucleic acid
CC	sequence of the GB virus-B clone and the HCV/GBV-B chimera may be used
CC	in the development of vaccines and therapeutics for HCV
XX	
SQ	Sequence 9595 BP; 1934 A; 2843 C; 2697 G; 2121 T; 0 U; 0 Other.

Query Match	88.9%	Score 1547	DB 4	Length 9595	
Best Local Similarity	93.9%	Prod. No. 0			
Matches 1610	Conservative	0	Mismatches 105	Indels 0	Gaps 0
QY	8	GCATGTCAATGTCTCTATACATATGACAGAGGCGCCCTGATCACACGTTGGCTTGCAGGAGAAA	67		
Db	7594	GCTGCTCAATGTCTCTATACGTGACACAGGCGCCCTGATCACGCCATGCGTTCGCGAGGAAA	7653		
QY	68	GCAAGCTGCCCATCAACAGGGCTGAGCAACTCTTTGCTGCTGCTACCACTAACCTGGTCTATT	127		
Db	7654	GTAAGCTGCCCATCAACCCGTTGAGCAACTCTTTGCTGCTGCTACCACTAACCTGGTCTACG	7713		
QY	128	CCACAACATCCCGCAGTGCAAAGCTCCGGCGAGAAAGGTACCTTTTGACAGACTGCAAG	187		
Db	7714	CCACAACATCCCGCAGCGCAAGCCTCCGCGAGAGAGGTACCTTTTGACAGATTGCAAG	7773		
QY	188	TCTTGACGATCATTTACCGGACGTGCTCAAGGAGATGAAGCGAAGGGCTCCACAGTGA	247		
Db	7774	TCTTGATGATCATTTACCGGACGTACTCAAGGAGATGAAGCGAAGGGCTCCACAGTTA	7833		
QY	248	AGGCTAAATCTGCTATCTGTAGAGAGCATGCAAGCTGACGCCCGCATTTGGCCAAAT	307		
Db	7834	AGGCTAAGCTTCTATCTATAGAGAGGCGCTGCAAGCTGACGCCCGCATTTGGCCAAAT	7893		
QY	308	CAAAATTTGGCTATGGGGCAAGGAGCTCCGAGAGCTATCCAGCAGGGCGGTTAACCA	367		
Db	7894	CAAAATTTGGCTATGGGGCAAGGAGCTCCGAGACCTATCCAGCAGGGCGGTTAACCA	7953		
QY	368	TCCGCTCCGTGTGGAAGGACTTGTCTGGAGGACACTCACACCAATTTAGAGCACCATCA	427		
Db	7954	TCCGCTCCGTGTGGAGGACTTGTCTGGAGGACACTGAAACCAATTTGACACCAATCA	8013		
QY	428	TGGCAAAAAATGAGGTTTTCTCGTCCAAACAGAGAAAGGAGGCGCGCAACCAAGCTCGCC	487		
Db	8014	TGGCAAAAAATGAGGTTTTCTCGTCCAAACAGAGAGGAGGCGCGCAAGCAGCTCGCC	8073		

QY	488	TCATCGTATTC	CCAGACCTGGAGTT	CGTGATATGCGAGNAGATGGCCCTTTACGACGTGG	547	Db	9154	CCACTTGTGGCAGATAC	CTCTTTAACTGGGCAGTAAAGACCAAGCTTTAAACTC	9211
Db	8074	TTATCGTATTC	CCAGACCTGGAGTT	CGTGATATGCGAGNAGATGGCCCTTTACGACGTGG	8133	QY	1628	TTCCGGCTCGCT	TCCCGGTGACCTTGTCCCGCTGGTTGCTGTGGCTACAGCGCGGGAG	1687
QY	548	TTTCCACTCTT	CTTCAGCGCGTGA	TGGGCTCCCTCATACGGATTCCAATACTCTCTTAAGC	607	Db	9214	TCCCGCGCGCT	TCCCGGTGACCTTGTCTGGCTGGTTACGCGGGGGAG	9273
Db	8134	TCTCCACCCCT	TCTTCAGCGCGTGA	TGGGCTCCCTCATACGGATTCCAATACTCTCTTAAGC	8193	QY	1688	ACATATATACAGC	CTGTCTCGTCCCGACCCCGC	1722
QY	608	AGCGGGTCGAGT	TCCTCGTGAATAC	CTGCAAGCAAGAAATGCGCTATGGGCTTCTCAT	667	Db	9274	ACATATATACAGC	CTGTCTCGTCCCGACCCCGC	9308
Db	8194	AGCGGGTCGAGT	TCCTCGTGAATAC	CTGCAAGCAAGAAATGCGCTATGGGCTTCTCAT	8253	RESULT 13				
QY	668	ATGACACCCGCT	GTTTGTGACTCA	ACCGTCACTGAGAAATGACATCCGTGTTGAGGAGTCAA	727	AAC86939 standard; DNA; 9595 BP.				
Db	8254	ATGACACCCGCT	GTTTGTGACTCA	ACCGTCACTGAGAAATGACATCCGTGTTGAGGAGTCAA	8313	XX	AAC86939;			
QY	728	TTTACCAATGTT	GAGCTTGGCCCCG	GAAGCTTAGACAGCGCCATTAAGTTCGCTCAACAGC	787	XX	AC	AAC86939;		
Db	8314	TTTACCAATGTT	GAGCTTGGCCCCG	GAAGCTTAGACAGCGCCATTAAGTTCGCTCAACAGC	8373	DT	02-APR-2001	(first entry)		
QY	788	GGCTCTATGTC	CGGGGTCCCATG	ACTTAACCTCAAAGGCGAGAAGCTTGCCTATCGCCGT	847	XX	02-APR-2001	(first entry)		
Db	8374	GGCTTTACATC	CGGGGTCCCATG	ACTTAACCTCAAAGGCGAGAAGCTTGCCTATCGCCGT	8433	DE	Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 1b.			
QY	848	GCCGCGCAGCG	CGTGTGACGACT	AGCTGCGGTAAATACCTCAATGCTTACTTTGAAGG	907	XX	Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;			
Db	8434	GCCGCGCAGT	GGCGTGTGACG	ACTAGCTGCGGTAAATACCTCAATGCTTACTTTGAAGG	8493	KW	HCV; vaccine; viral inhibitor; antiviral; ss.			
QY	908	CGCTGCACCT	GTCCAGCTGCCA	AGCTCCAGGATGCTCGTGAATGAGACG	967	XX	Hepatitis C virus.			
Db	8494	CCACTGCACCT	GTCCAGCTGCCA	AGCTCCAGGATGCTCGTGAATGAGACG	8553	XX	Key	Location/Qualifiers		
QY	968	ACCTTGTCTG	TATCTGTGAAG	CGCGGAAACCCAGAGGACCGCGCAAGCTTACGAGTCT	1027	FT	CDS	342..9374		
Db	8554	ACCTTGTCTG	TATCTGTGAAG	CGCGGAAACCCAGAGGATCGGGGGCCCTACGAGCT	8613	FT	/*tag= a			
QY	1028	TACGAGGCTAT	GACTAGGTACT	TGCCCCCTTGGGGACCGCCCCAACCGGAATAG	1087	PN	WO200075352-A2.			
Db	8614	TACGAGGCTAT	GACTAGGTACT	TGCCCCCTTGGGGATCGGGGGCCCTACGAGTACG	8673	PD	14-DEC-2000.			
QY	1088	ACTTGGAGCT	GATACATCGT	GTCTCCCAATGTGTGCGTGCACACGATGCTGCA	1147	XX	02-JUN-2000; 2000WO-US015527.			
Db	8674	ACCTGAGCT	GATACATCAT	GTTCCTCCAATGTGTGAGTGCACACGATGCTGCA	8733	XX	04-JUN-1999; 99US-0137817P.			
QY	1148	AAAGGTGACT	ACTCACCGGTG	ACCCGCTTGGCGGGCTGGTGGGAGA	1207	PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.			
Db	8734	AAAGGTGATA	CTCACCGGTGAC	CCACCGCTTGGCGGGCTGGTGGGAGA	8793	XX	Nam J, Bukh J, Emerson SU, Purcell RH;			
QY	1208	CAGCTAGGCA	CACTCCAGTCA	ACTCTCGCTAGGCAACATCATGTATGGCCCACTT	1267	XX	WPI; 2001-071081/08.			
Db	8794	CAGCTAGACA	CACTCCATCACT	CTTGGCTAGGCAATATCATCATGTATGGCCCACTT	8853	XX	P-PSDB; AAB31170.			
QY	1268	TGTGGGCAAG	GATGTTGATG	ACTCTTCTCATCTTCTAGCCAGGAGCAAC	1327	XX	New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome			
Db	8954	TATGGGCAAG	GATGTTGATG	ACTCTTCTCATCTTCTAGCTCAAGAGCAAC	8913	XX	in which the (non-)structural region has been replaced by hepatitis C			
QY	1328	TTGAAAAGCC	CTTGATGACT	ACGAGCTTGTACTCCATGAGCCACTTGACC	1387	XX	virus (HCV) genome useful for treating or preventing HCV signs and			
Db	8914	TTGAAAAGCC	CTTGATGACT	ACGAGCTTGTACTCCATGAGCCACTTGACC	8973	XX	symptoms.			
QY	1388	TACCTGAGAT	CATGTGAGCT	CACTTCTTACGCAATTTTCACTCCAAGTACTCTC	1447	PS	Disclosure; Fig 4A-F; 97pp; English.			
Db	8974	TACCTGAGAT	CATGTGAGCT	CACTTCTTACGCAATTTTCACTCCAAGTACTCTC	9033	XX	The specification describes a nucleic acid comprising a chimeric virus			
QY	1448	CAGGTGAGAT	CAATAGGGTGT	TCATGCTTCAGGAAAGCTTGGGGTACCAACCTTGGAG	1507	CC	genome, specifically bovine viral diarrhoea virus (BVDV) genome in which			
Db	9034	CAGGTGAGAT	CAATAGGGTGT	TCATGCTTCAGGAAAGCTTGGGGTACCAACCTTGGAA	9093	CC	the (non-)structural region has been replaced by the (non-)structural			
QY	1508	TCTGGAGACA	TCGGGCCAG	AGTGTCCGGCTAAGTTACTGTCTCCAGGGGGAGGGCCG	1567	CC	region of a hepatitis C virus (HCV) genome. The nucleic acids comprising			
Db	9094	CCTGGAGACA	TCGGGCCAG	AGTGTCCGGCTAAGTTACTGTCTCCAGGGGGAGGGCCG	9153	CC	the chimeric virus and the chimeric virus are useful for identifying cell			
QY	1627	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	lines capable of supporting the replication of these chimeric viruses, in			
Db	9153	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	screening for neutralizing antibodies to HCV of different genotypes, in			
QY	1627	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	the production of HCV-BVDV virions, for the development of inactivated or			
Db	9153	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the			
QY	1627	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	molecular properties of HCV indirectly in vitro, and in identifying			
Db	9153	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	inhibitors of viral enzyme activity which would be useful as antiviral			
QY	1627	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	agents. Formulations or compositions comprising the chimeric viruses may			
Db	9153	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	be used to treat or prevent the signs and symptoms of HCV. The present			
QY	1627	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	sequence represents a HCV clone, which is used to construct chimeric			
Db	9153	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	CC	nucleic acids of the invention			
QY	1627	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	XX	Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;			
Db	9153	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	XX	Query Match 88.9%; Score 1547; DB 4; Length 9595;			
QY	1627	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	XX	Best Local Similarity 93.9%; Pred. No. 0;			
Db	9153	CCATTTGGGCA	AGTACTCTT	CAACTGGGAGTGAAGGACCAAGCTTAAACTACTCCAA	1627	XX				

Db	9154	CCACTTGTGGCAGATACCTCTTTAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAA	9213
Qy	1628	TTCCGGCTCGTCCCGGCTGGACTTGTCCGGCTGGTTGTTGCTGCTACAGCGGGGAG	1687
Db	9214	TCCCGGCGCGTCCCAGCTGGACTTGTCTGGCTGGTTGCTGCTGCTTACAGCGGGGAG	9273
Qy	1688	ACATATATCACAGCTGTCTCGTGCCCGACCCCGC	1722
Db	9274	ACATATATCACAGCTGTCTCGTGCCCGACCCCGC	9308
RESULT 13			
AAC86939			
ID	AAC86939 standard; DNA; 9595 BP.		
XX	AAC86939;		
AC			
XX			
DT	02-APR-2001 (first entry)		
XX			
DE	Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 1b.		
XX			
XX	Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;		
KW	HCV; vaccine; viral inhibitor; antiviral; ss.		
XX			
OS	Hepatitis C virus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	342..9374	
FT		/*tag= a	
XX			
PN	WO200075352-A2.		
XX	14-DEC-2000.		
XX	02-JUN-2000; 2000WO-US015527.		
XX	04-JUN-1999; 99US-0137817P.		
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
PA			
XX	Nam J, Bukh J, Emerson SU, Purcell RH;		
PI			
XX	WPI; 2001-071081/08.		
DR	P-PSDB; AAB31170.		
DR			
XX	New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome		
PT	in which the (non-)structural region has been replaced by hepatitis C		
PT	virus (HCV) genome useful for treating or preventing HCV signs and		
PT	symptoms.		
XX			
PS	Disclosure; Fig 4A-F; 97pp; English.		
XX			
CC	The specification describes a nucleic acid comprising a chimeric virus		
CC	genome, specifically bovine viral diarrhoea virus (BVDV) genome in which		
CC	the (non-)structural region has been replaced by the (non-)structural		
CC	region of a hepatitis C virus (HCV) genome. The nucleic acids comprising		
CC	the chimeric virus and the chimeric virus are useful for identifying cell		
CC	lines capable of supporting the replication of these chimeric viruses, in		
CC	screening for neutralizing antibodies to HCV of different genotypes, in		
CC	the production of HCV-BVDV virions, for the development of inactivated or		
CC	attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the		
CC	molecular properties of HCV indirectly in vitro, and in identifying		
CC	inhibitors of viral enzyme activity which would be useful as antiviral		
CC	agents. Formulations or compositions comprising the chimeric viruses may		
CC	be used to treat or prevent the signs and symptoms of HCV. The present		
CC	sequence represents a HCV clone, which is used to construct chimeric		
CC	nucleic acids of the invention		
XX			
SQ	Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;		
Query Match 88.9%; Score 1547; DB 4; Length 9595;			
Best Local Similarity 93.9%; Pred No. 0;			

PR 15-NOV-2002; 2002GB-00026722.
XX (GLAX) GLAXO GROUP LTD.
PA Brett S, Hamblin PA, Ogilvie L;
XX P-PSDB; ADO36227.
DR WPI; 2004-420613/39.
XX PT New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT the polypeptide sequences of the HCV core and at least one other HCV
PT protein, for use in medicine, particularly for manufacturing a medicament
PT for treating HCV.
XX Disclosure; Fig 1; 78pp; English.
PS
XX The invention describes a polynucleotide vaccine comprising a
CC polynucleotide sequence (S1) encoding the Hepatitis C virus (HCV) Core
CC protein and at least 1 other HCV protein, and causes expression of the
CC proteins in cells (in which (S1) has been mutated or positioned relative
CC to the polynucleotide sequence encoding the other HCV protein, so that
CC the negative effect of the Core protein on expression of the other HCV
CC protein is reduced). Also described are: a method of preventing or
CC treating an HCV infection in a mammal, comprising administering the
CC vaccine cited above to a mammal; and a method of vaccination of an
CC individual, comprising taking a polynucleotide vaccine as cited above,
CC coating the polynucleotide onto gold beads and delivering the gold beads
CC into the skin. HCV nucleic acids, polypeptides, host cells, vectors and
CC antibodies used in the methods, are also disclosed. The polynucleotide
CC vaccine is useful in the manufacture of a medicament for the treatment of
CC HCV. This sequence encodes the wild type HCV polypeptide.
XX
SQ Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;

Query Match 88.9%; Score 1547; DB 12; Length 9595;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1610; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY	8	GCATGTCAATGCTCTATACATGACAGGCGCCCTGTATACACCGTGGCTGCGGAGGAA	67
DB	7594	GCTGCTCAATGTCTTATACGTGACAGGCGCCCTGATCAACGCGATGCGCTGCGGAGGAA	7653
QY	68	GCAAGCTGCCATCAACGCGCTGAGCACTCTTGTGCTGCTCAACCAATACCTGGTCTATT	127
DB	7654	GTAAGCTGCCCATCAACCCGTTGAGCAACTCTTGTGCTGCTCAACCAATACCTGGTCTATT	7713
QY	128	CCACACATCCCGCAGTGCAGGCTCGGCGAGAGAGGTTCACCTTTGACAGACTGCAAG	187
DB	7714	CCACACATCCCGCAGGCGGAGGCTCGGCGAGAGAGGTTCACCTTTGACAGATTGCAAG	7773
QY	188	TCCTGGACGATCATTAACCGGAGCTGCTCAAGGAGATGAAGGAGGCGTCCACAGTGA	247
DB	7774	TCCTGGATGATCATTAACCGGAGCTCTCAAGGAGATGAAGGAGGCGTCCACAGTTA	7833
QY	248	AGGCTAACTGTCTATCTGTAGAGAGAGCATGCAAGCTGACGCGCCCGCATTCGGCCAAAT	307
DB	7834	AGGCTAAGCTTCTATCTATAGAGAGGCGCTGCAAGCTGACGCGCCCGCATTCGGCCAAAT	7893
QY	308	CCAAATTTGGCTATGGGCGAAAGAGCTCGGAGCCCTATCCAGCAGGCGGTTTAAACACA	367
DB	7894	CCAAATTTGGCTATGGGCGAAAGAGCTCGGAGCCCTATCCAGCAGGCGGTTTAAACACA	7953
QY	368	TCGCTCCGTGTGAAGACTTGTCTGGAGGACACTGACACACCAATTCAGACCACTATCA	427
DB	7954	TCCGCTCCGTGTGGAGAGACTTGTCTGGAGGACACTGAAACCAATTCAGACCACTATCA	8013
QY	428	TGGCAAAAATAGAGTTTCTGCGTCCAAACAGAGAGAGGCGCGCAAAACAGAGTGGCC	487
DB	8014	TGGCAAAAATAGAGTTTCTGCGTCCAAACAGAGAGAGGCGCGCAAAACAGAGTGGCC	8073
QY	488	TCATCGTATNTCCAGACCTGGGAGTTGCTGTATCGGAGAGATGGCCCTTTACGACGTGG	547
DB	8074	TTATCGTATNTCCAGACCTGGGAGTTGCTGTATCGGAGAGATGGCCCTTTACGACGTGG	8133

QY	548	TTTCCACTCTTCTCCTCAGGCGCTGATGGGCTCTCTCATACGGAATCCAATATCTCTCTAAGC	607
DB	8134	TTCTCCACCCCTTCTCCTCAGGCGCTGATGGGCTCTCTCATACGGAATCCAATATCTCTCTAAGC	8193
QY	608	AGCGGTCGAGTTTCTCTGGTGAATACCTGGAAAGCAAGAAATGCCCCCTATGGGTTCTCAT	667
DB	8194	AGCGGTCGAGTTTCTCTGGTGAATACCTGGAAATCAAGAAATGCCCCCTATGGGTTCTCAT	8253
QY	668	ATGACACCCGCTCTTTTGTACTCAACGCTCACTGAGATGACATCCGTTGTGAGGACTCAA	727
DB	8254	ATGACACCCGCTCTTTTGTACTCAACGCTCACTGAGAGTGAATTCGTTGTGAGGAGTCAA	8313
QY	728	TTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCGCATAAAGTCTCCTCACAGAGC	787
DB	8314	TTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCGCATAAAGTCTCCTCACAGAGC	8373
QY	788	GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGGCAGAACTCGGGCTATCGCGGCT	847
DB	8374	GGCTTTTACATCGGGGTCCTCTGACTCACTCAAAAGGGCAGAACTCGGGTTATCGCGGCT	8433
QY	848	GCCGCGGAGCGGCTGCTGACCACTAGCTGCGGTAAATACCTCAATGCTACTTGAAGG	907
DB	8434	GCCGCGCAAGTGGGCTGCTGACGACTAGCTGCGGTAAATACCTCAATGCTACTTGAAGG	8493
QY	908	CCGCTGAGGCTGTCGAGCTGCCAAGCTCCAGGACTGACAGATGCTCGTGAATGGAGAGC	967
DB	8494	CCACTGCGAGCTGTCGAGCTGCCAAGCTCCAGGACTGACAGATGCTCGTGAATGGAGAGC	8553
QY	968	ACCTTGTCTTATCTGTGAAAGCGCGGAAACCCAAAGAGGAGCGGGCAAGCTTACAGTCT	1027
DB	8554	ACCTTGTCTTATCTGTGAAAGCGCGGAAACCCAAAGAGGAGTGGCGGCGCTTACAGACCT	8613
QY	1028	TCAGGAGGCTATGACTAGGTACTCTGCCCCCCCCCTGGGGAGCCGCCCAACCGGAATACG	1087
DB	8614	TCAGGAGGCTATGACTAGGTATTTCCGCCCCCCCCCGGGATCCGCCCAACCAAGATACG	8673
QY	1088	ACTTGGAGCTGATAAATCTGTTTCTTCCAAATGTTGTCGGTGGCAACAGATGATCTGGCA	1147
DB	8674	ACTTGGAGCTGATAAATCTGTTTCTTCCAAATGTTGTCAGTCGGCGCAGATGATCTGGCA	8733
QY	1148	AAAGGTTGTACTACTCTACCGTGAACCCCAAGTCCCTTGGCGGGCTGCGTGGGAGA	120
DB	8734	AAAGGTTGTACTACTCTACCGTGAACCCCAAGTCCCTTGGCGGGCTGCGTGGGAGA	879
QY	1208	CAGCTAGGACACTCCAGTCAACTCTGGCTAGGCAACATCATATGATATGCGGCCACTT	126
DB	8794	CAGCTAGGACACTCCAGTCAACTCTGGCTAGGCAACATCATATGATATGCGGCCACTT	885
QY	1268	TGTGGGCAAGATGATTTCTGATGATCTCTTCTTCTCATCTTCTTACGCCAGGAGAAC	132
DB	8854	TATGGGCAAGATGATTTCTGATGATCTTGGCTAGGCAATATATCATATGATGGGCCACC	891
QY	1328	TTGAAAAAGCCCTGATTTGTGATGATCTACGGGCTTGTACTCCATGAGCCACTTGGCC	138
DB	8914	TTGAAAAAGCCCTGATTTGTGATGATCTTCTTCTCATCTTCTTCTCATTTGAGCCACTT	897
QY	1388	TACCTCAGATCATTTGAACGACTCCATGCTTACGCGCATTTTACCTCCATAGTACTCTC	144
DB	8974	TACCTCAGATCATTTGAACGACTCCATGCTTACGCGCATTTTACCTCCATAGTACTCTC	903
QY	1448	CAGGTCAGATCAATAGGTTGGCTTTCATGCTCAGGAAAGTTGGGGTACACCTTTGGAG	150
DB	9034	CAGGTCAGATCAATAGGTTGGCTTTCATGCTCAGGAAAGTTGGGGTACACCTTTGGAG	909
QY	1508	TCTGGAGACATCGGGCCAGAGTGTCCGCGCTTAAGTTACTGTCCAGGGGGGAGGCGCG	156
DB	9094	CCTGGAGACATCGGGCCAGAGTGTCCGCGCTTAAGTTACTGTCCAGGGGGGAGGCGCG	915
QY	1568	CCATTTGTGGCAAGTACTCTTCAACTGGGAGTAAAGGACCAAGCTTAACTCACTTCCAA	162
DB	9154	CCACTTTGGCAGATACCTCTTTTAACTGGGAGTAAAGGACCAAGCTTAACTCACTTCCAA	921

```
QY 1628 TTCCGGCTGGTCCCGGCTGGAGCTTGTCCGGCTGGTTCGTTGCTGCTACAGCGGGGAG 1687
Db 9214 TCCCGCCGCGTCCAGCTGGACTTGTCTGGCTGGTTCGTTGCTGCTTACAGCGGGGAG 9273
QY 1688 ACATATATCACAGCCTGCTCGTCCGCGACCCCGC 1722
Db 9274 ACATATATCACAGCCTGCTCGTGGCCGACCCCGC 9308

RESULT 15
AD079396
ID AD079396 standard; cDNA; 5955 BP.
XX
AC AD079396;
XX
DT 26-AUG-2004 (first entry)
XX
DE Hepatitis C virus J4L6 genome wild-type cDNA.
XX
KW HCV; polyprotein; vaccine; DNA immunisation; hepatotropic; virucide;
KW gene; ss.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 342..9374
FT /tag= a
FT /product= "HCV polyprotein"
XX
PN WO2004046176-A1.
XX
PD 03-JUN-2004.
XX
PF 13-NOV-2003; 2003WO-EP012830.
XX
PR 15-NOV-2002; 2002GB-00026722.
XX
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Brett S, Hamblin PA, Ogilvie L;
XX
WPI; 2004-420614/39.
DR P-PSDB; AD079401.
DR GENBANK; AF054247.
XX
XX
XX New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT the polypeptide sequences of the HCV core, NS3, NS4B and NS5B proteins,
PT for use in medicine, in particular for manufacturing a medicament for the
PT treatment of HCV.
XX
PS Disclosure; Fig 1; 79pp; English.
XX
XX
XX The present sequence is the hepatitis C virus (HCV) J4L6 genome wild-type
CC cDNA sequence encoding the HCV polyprotein. HCV DNA vaccines of the
CC invention comprise a polynucleotide that encodes the HCV proteins Core,
CC NS3, NS4B and NS5B, and does not encode the NS4A and/or NS5A proteins.
CC Preferably, the codon usage of the polynucleotide sequence resembles that
CC of highly expressed human genes. The polynucleotides may encode
CC individual proteins or fusion proteins. Preferred fusions include double
CC fusions between NS4B and NS5B and between Core and NS3. The vaccines are
CC useful for the treatment or prevention of an HCV infection.
XX
SQ Sequence 5955 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;

Query Match 88.9%; Score 1547; DB 12; Length 9595;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1610; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 8 GCATGTCAATGCTCTATACATGGACAGCGCCCTGTATCACACCGTGGCTGGGAGGAAA 67
Db 7594 GCTGCTCAATGCTCTATACATGGACAGCGCCCTGTATCACACCGTGGCTGGGAGGAAA 7653
```

```
QY 68 GCAAGCTGCCCATCAACGCGCTGAGCAACTCTTGTCTGCTGCTACCATCAACCTGGTCTATT 127
Db 7654 GTAAGCTGCCCATCAACCCGTTGAGCAACTCTTGTCTGCTGCTACCAACCATGGTCTAGC 7713
QY 128 CCACAACATCCCGCAGTGCAAGCCCTGCGGCGAGAAGAGTCAACCTTTTGACAGACTGCAAG 187
Db 7714 CCACAACATCCCGCAGCGCAAGCCCTCGGCGAGAAGAGTCAACCTTTTGACAGATTGCAAG 7773
QY 188 TCCTGGACGATCATATTACGGGACGTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTGA 247
Db 7774 TCCTGGATGATCATTTACCGGACGCTACTCAAGGAGATGAAGCGGAGCGTCCACAGTTA 7833
QY 248 AGGCTAAACTGCTATCTGTAGAAGAGCATGCAAGCTGCAAGCCCGCCGCAATTCGGCCAAAT 307
Db 7834 AGGCTAAGCTTCTATCTATAGAGGAGGCTGCAAGCTGCAAGCCCGCCGCAATTCGGCCAAAT 7893
QY 308 CCAAAATTTGGCTATGGGCGAAAGACGCTCCGGAGCTTATCCAGCAGGCGCTTTAAACACA 367
Db 7894 CCAAAATTTGGCTATGGGCGAAAGACGCTCCGGAGCTTATCCAGCAGGCGCTTTAAACACA 7953
QY 368 TCCGCTCGTGTGGAAGGACTTGTGAGGAGCACTGACACACCAATTTGACACCACTCA 427
Db 7954 TCCGCTCGTGTGAGGAGCTTGTGGAAGACACTGAAACACCAATTTGACACCACTCA 8013
QY 428 TGGCAAAATATGAGGTTTCTGCTGCCAACAGAGAAAGAGGCGCGAAACCAAGCTCGCC 487
Db 8014 TGGCAAAATATGAGGTTTCTGCTGCCAACAGAGAAAGAGGCGCGAAACCAAGCTCGCC 8073
QY 488 TCATCGTATTTCCAGACCTCGGAGTTCGTGTATCGGAGAGATGGCCCTTTTACACGTGG 547
Db 8074 TTATCGTATTTCCAGACCTCGGAGTTCGTGTATCGGAGAGATGGCCCTTTTACACGTGG 8133
QY 548 TTTCCTACTCTTCTCAGGCGCTGATGGGCTCCTCATACGGATTCCAATACTCTCTTAAGC 607
Db 8134 TCTCCACCTTCTCAGGCGCTGATGGGCTCCTCATACGGATTCCAATACTCTCTCCCAAGC 8193
QY 608 AGCGGTCGAGTTCTCTGGTGAATACTCTGGAAGCAAAAGAAATGCCCTATGGGCTTCTCAT 667
Db 8194 AGCGGTCGAGTTCTCTGGTGAATACTCTGGAAGCAAAAGAAATGCCCTATGGGCTTCTCAT 8253
QY 668 ATGACACCCGCTGTTTGACTCAACGCTCACTGAGAAATGACATCCGTTGTAGAGAGTCAA 727
Db 8254 ATGACACCCGCTGTTTGACTCAACGCTCACTGAGAGTGAATTCGTTGTAGAGAGTCAA 8313
QY 728 TTTTACCAATGTTGTGACTTCGCGCCCGAAGCTAGACAGGCGCATAGGCTCGCTCACAGC 787
Db 8314 TTTTACCAATGTTGTGACTTCGCGCCCGAAGCTAGACAGGCGCATAGGCTCGCTCACAGC 8373
QY 788 GGCTCTATGTCGGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGCGGCTATTCGCCGCT 847
Db 8374 GGCTTTATATCGGGGGTCCCTGACTAACTCAAAAGGCGAGAACTGCGGCTATTCGCCGCT 8433
QY 848 GCCCGGAGCGCGCTGCTGACGACTAGTTCGGTAAATACCTTCACATGCTACTTTGAAG 907
Db 8434 GCCCGCAAGTGGCGTCTGACGACTAGTTCGGGTAATACCTTCACATGTTACTTTGAAG 8493
QY 908 CCGCTCGAGCGCTCGAGCTGCCAAGCTCCAGACTCCAGCATGCTCGTGAATGGAGAGC 967
Db 8494 CCACTCGAGCGCTCGAGCTGCCAAGCTCCAGACTCCAGCATGCTCGTGAATGGAGAGC 8553
QY 968 ACCTTGTGTTATCTGTGAAAGCGGGGAACCCAGAGAGCGCGGCAAGCTTACGAGTCT 1027
Db 8554 ACCTTGTGTTATCTGTGAAAGCGGGGAACCCAGAGAGATGCGGCGGCTTACGAGCT 8613
QY 1028 TCACGGAGGCTATGACTAGTACTCTGCCCGCCCTCGGGACCGCCCGCAACCGGAATACG 1087
Db 8614 TCACGGAGGCTATGACTAGTATTCGCCCGCCCGCCCGGGATCCGCCCGCAACAGAAATACG 8673
QY 1088 ACTTGAGCTGATPAAATCGTGTTCCTCCAAATGTTGCGGTGCGCACAGCATGCTTGGA 1147
Db 8674 ACCTGAGCTGATPAAATCATGTTCTCTCCAAATGTTGTCAGTTCGCGCAGCATGCTTGGA 8733
QY 1148 AAAGGGTGTTACTACCTACCCGCTGACCCCGTCCCTTTCGCGGGGCTGCGTGGGAGA 1207
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 14:01:52 ; Search time 1226.79 Seconds
(without alignments)
9793.114 Million cell updates/sec

Title: US-10-712-479-3

Perfect score: 1740

Sequence: 1 atgctagcatgtcatgtc.....gccatcaccatcaccatcac 1740

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1740	100.0	1740	20	US-10-712-479-3
2	1722	99.0	1803	20	US-10-712-479-1
3	1555.6	89.4	1884	9	US-09-638-386-11
4	1543.8	88.7	5955	20	US-10-492-178-5
5	1537.4	88.4	5965	20	US-10-492-178-2
6	1537.4	88.4	37090	20	US-10-492-178-4
7	1527.8	87.8	6189	15	US-10-259-275-41

8	1527.8	87.8	6189	24	US-11-006-313-41	Sequence 41, Appl
9	1527.8	87.8	7989	17	US-10-434-842-16	Sequence 16, Appl
10	1527.8	87.8	7989	19	US-10-434-842-16	Sequence 1, Appl
11	1527.8	87.8	7989	21	US-10-897-648-17	Sequence 17, Appl
12	1527.8	87.8	7992	13	US-10-005-469-1	Sequence 2, Appl
13	1527.8	87.8	7992	13	US-10-005-469-2	Sequence 1, Appl
14	1527.8	87.8	7992	13	US-10-005-469-4	Sequence 4, Appl
15	1527.8	87.8	7992	13	US-10-005-469-5	Sequence 5, Appl
16	1527.8	87.8	7992	13	US-10-005-469-6	Sequence 6, Appl
17	1527.8	87.8	7992	17	US-10-434-842-1	Sequence 2, Appl
18	1527.8	87.8	7992	17	US-10-434-842-2	Sequence 1, Appl
19	1527.8	87.8	7992	17	US-10-434-842-4	Sequence 4, Appl
20	1527.8	87.8	7992	17	US-10-434-842-5	Sequence 5, Appl
21	1527.8	87.8	7992	17	US-10-434-842-6	Sequence 6, Appl
22	1527.8	87.8	7992	17	US-10-434-842-15	Sequence 15, Appl
23	1527.8	87.8	7995	13	US-10-005-469-3	Sequence 3, Appl
24	1527.8	87.8	7995	17	US-10-434-842-3	Sequence 3, Appl
25	1527.8	87.8	8732	22	US-10-510-912-1	Sequence 1, Appl
26	1527.8	87.8	9605	18	US-10-467-000-2	Sequence 2, Appl
27	1527.8	87.8	10690	14	US-10-125-940-1	Sequence 1, Appl
28	1527.8	87.8	10690	16	US-10-125-920-1	Sequence 1, Appl
29	1527.8	87.8	10690	18	US-10-467-000-3	Sequence 3, Appl
30	1527.8	87.8	12305	20	US-10-422-323A-2	Sequence 2, Appl
31	1527.8	87.8	12315	20	US-10-422-323A-1	Sequence 1, Appl
32	1527.6	87.8	1772	19	US-10-384-339C-132	Sequence 132, App
33	1526.2	87.7	7992	17	US-10-434-842-17	Sequence 17, Appl
34	1526.2	87.7	8638	13	US-10-029-907-6	Sequence 6, Appl
35	1526.2	87.7	8638	13	US-10-029-907-7	Sequence 7, Appl
36	1526.2	87.7	8638	13	US-10-029-907-24	Sequence 24, Appl
37	1526.2	87.7	8638	13	US-10-029-907-25	Sequence 25, Appl
38	1526.2	87.7	8638	15	US-10-309-561-6	Sequence 6, Appl
39	1526.2	87.7	8638	15	US-10-309-561-7	Sequence 7, Appl
40	1526.2	87.7	8638	15	US-10-309-561-24	Sequence 24, Appl
41	1526.2	87.7	8638	19	US-10-309-561-25	Sequence 25, Appl
42	1526.2	87.7	8638	19	US-10-789-355-6	Sequence 6, Appl
43	1526.2	87.7	8638	19	US-10-789-355-7	Sequence 7, Appl
44	1526.2	87.7	8638	19	US-10-789-355-24	Sequence 24, Appl
45	1526.2	87.7	8638	19	US-10-789-355-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-10-712-479-3
; Sequence 3, Application US/10712479
; Publication No. US20040209283A1
; GENERAL INFORMATION:
; APPLICANT: Yagi, Yoshihiko
; APPLICANT: Sheets, Michael P.
; APPLICANT: Wells, Peter A.
; APPLICANT: Shelly, John A.
; APPLICANT: Foorman, Roger A.
; APPLICANT: Epps, Dennis E.
; TITLE OF INVENTION: A Continuous-Read Assay for the Detection of De Novo HCV RNA Poly
; FILE OF INVENTION: Activity
; CURRENT APPLICATION NUMBER: US/10712,479
; CURRENT FILING DATE: 2003-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: C-terminally truncated HCV NS5B polymerase (C delta 21 NSSB)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1740)
; OTHER INFORMATION:
US-10-712-479-3

; OTHER INFORMATION:

US-10-712-479-1

Query Match 99.0%; Score 1722; DB 20; Length 1803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTAGCATGTCATATCATGACAGCGCCCTGATCACACCGTGGCGTGG 60
DB 1 ATGGCTAGCATGTCATATCATGACAGCGCCCTGATCACACCGTGGCGTGG 60

QY 61 GAGGAAGCAAGCTGCCATCAACGCGCTGAGCAACTCTTGGTGGTCAACATACCTG 120
DB 61 GAGGAAGCAAGCTGCCATCAACGCGCTGAGCAACTCTTGGTGGTCAACATACCTG 120

QY 121 GTCTATTCCACAACTCCCGCAGTGCAGCCTGCGCAGAGAGTCACTTTGACAGA 180
DB 121 GTCTATTCCACAACTCCCGCAGTGCAGCCTGCGCAGAGAGTCACTTTGACAGA 180

QY 181 GTGCAAGTCTGAGCATCATTTACCGGAGCGTGTCAAGGAGATGAAGCGAAGCGTCC 240
DB 181 GTGCAAGTCTGAGCATCATTTACCGGAGCGTGTCAAGGAGATGAAGCGAAGCGTCC 240

QY 241 ACAGTGAAGGCTAACTGCTATCTGTAGAAAGCATGCAAGCTGACGCCCGCATTCG 300
DB 241 ACAGTGAAGGCTAACTGCTATCTGTAGAAAGCATGCAAGCTGACGCCCGCATTCG 300

QY 301 GCCAAATCCAAATTTGGCTATGGGCAAGAGCGTCCGAGCCTATCAGAGGCGCGTT 360
DB 301 GCCAAATCCAAATTTGGCTATGGGCAAGAGCGTCCGAGCCTATCAGAGGCGCGTT 360

QY 361 AACCAATCCCGTCCGTTGGAAGACTTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 420
DB 361 AACCAATCCCGTCCGTTGGAAGACTTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 420

QY 421 ACCATCATGGCAAAATAGGTTTCTGCGTCCAAACAGAGAGAGAGCGCGCAAAACA 480
DB 421 ACCATCATGGCAAAATAGGTTTCTGCGTCCAAACAGAGAGAGAGCGCGCAAAACA 480

QY 481 GCTGCGCTCATGTAATCCGAGACTGCGGAGTGTGATGCGAAGATGCGCCTTTAC 540
DB 481 GCTGCGCTCATGTAATCCGAGACTGCGGAGTGTGATGCGAAGATGCGCCTTTAC 540

QY 541 GAGTGGTTCACACTTCTCAGCGCGTGTGATGCGGCTCTCATAGGATTCGAATCTCT 600
DB 541 GAGTGGTTCACACTTCTCAGCGCGTGTGATGCGGCTCTCATAGGATTCGAATCTCT 600

QY 601 CCTAAGCAGCGGTGAGTTCCTGCTGAATACCTGGAAGCAAGAGATGCCCTATGGC 660
DB 601 CCTAAGCAGCGGTGAGTTCCTGCTGAATACCTGGAAGCAAGAGATGCCCTATGGC 660

QY 661 TTCTCATATGACACCGCTGTTTGAATCAACGGTCACTGAGAAATGACATCCGTTGAG 720
DB 661 TTCTCATATGACACCGCTGTTTGAATCAACGGTCACTGAGAAATGACATCCGTTGAG 720

QY 721 GAGTCAATTTACCAATGTTGACTGTTGSCCCCGAAGCTAGACAGGCGCAATAGGTCGCTC 780
DB 721 GAGTCAATTTACCAATGTTGACTGTTGSCCCCGAAGCTAGACAGGCGCAATAGGTCGCTC 780

QY 781 ACAGAGCGGCTCATGTCGGGGTCCCATGACTAATCTCCAAAGGCGAAGCTCGGCTAT 840
DB 781 ACAGAGCGGCTCATGTCGGGGTCCCATGACTAATCTCCAAAGGCGAAGCTCGGCTAT 840

QY 841 CGCCGGTCCCGGAGCGGCTGCTGACGACTAGCTGCGGTATACCTTCACATGCTAC 900
DB 841 CGCCGGTCCCGGAGCGGCTGCTGACGACTAGCTGCGGTATACCTTCACATGCTAC 900

QY 901 TTGAAGCGCGTGCAGGCTGTGAGCTGCCAAGCTCCAGGACTGCAAGATGCTCGTGAAT 960
DB 901 TTGAAGCGCGTGCAGGCTGTGAGCTGCCAAGCTCCAGGACTGCAAGATGCTCGTGAAT 960

QY 961 GGAGACAGCTTGTGCTTATCTGTGAAAGCGCGGAAACCCAAAGAGGACGCGCAAGCTA 1020
DB 961 GGAGACAGCTTGTGCTTATCTGTGAAAGCGCGGAAACCCAAAGAGGACGCGCAAGCTA 1020

RESULT 3

US-09-838-386-11

; Sequence 11, Application US/09838386

; Patent No. US20010055756A1

; GENERAL INFORMATION:

; APPLICANT: Pellerin, Charles

; APPLICANT: Kukoij, George

; TITLE OF INVENTION: Internal De No. US20010055756A1o Initiation Sites of the HCV NS5S

; TITLE OF INVENTION: Theroof

; FILE REFERENCE: 1011.2180001

; CURRENT APPLICATION NUMBER: US/09/838,386

; CURRENT FILING DATE: 2001-04-20

; PRIOR APPLICATION NUMBER: US 60/198,793

; PRIOR FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 11

; LENGTH: 1884

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: misc_feature

OTHER INFORMATION: Htra5B polymerase
; NAME/KEY: CDS
; LOCATION: (1)..(1881)
US-09-838-386-11

Query Match 89.4%; Score 1555.6; DB 9; Length 1884;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy	8	GCATGTCATATGCTCTATATGAGCAGCGCGCCCTGATCACACCGTGGCTCGCGAGGAAA	67
Db	104	GCTGCTGATGTCCTACACATGAGCAGCGCGCCCTGATCACGCCATGGCGCGGAGGAAA	163
Qy	68	GCAAGCTGCCATCAACGGCTGAGCAACTCTTGGCTGGCTGACCATCAACCTGCTCTATT	127
Db	164	GCCAGCTGCCATCAACGGCTGAGCAACTCTTGGTGGCTGATCGCAACATGGTCTATT	223
Qy	128	CCACAACTCCCGCAGTGCAGCCCTGGCGAGAGAGGTCACTTTGACAGACTGCAAG	187
Db	224	CCACAACTCCCGCAGCGCGCCCTGGCGAGAGAGGTCACTTTGACAGACTGCAAG	283
Qy	188	TCCTGACGATCATTTACCGGACGCTGCTCAAGGAGATGAAGCGGAGCGCTCCACAGTGA	247
Db	284	TCCTGACGACCACTACCGGACGCTGCTCAAGGAGATGAAGCGGAGCGCTCCACAGTTA	343
Qy	248	AGGCTAACTGCTATCTGTGAGAGAGCATGCAAGCTGACGCGCCCGCATTTGGGCCAAT	307
Db	344	AGGCCAACTACTATCAGTGAAGAGCGCTGTAAGCTGACGCGCCCGCATTTGGGCCAAT	403
Qy	308	CCAAATTTGGCTATGGGCGAAAGAGCGTCCGAGCCTATCCAGCAGGCGCGTTACCAACA	367
Db	404	CCAAGTTTGGCTATGGGCGAAAGAGCGTCCGAGCCTATCCAGCAGGCGCGTTGACCAACA	463
Qy	368	TCGCTCCGCTGGAGAGACTTGTGGAGGACACTGACACACCAATTCAGACCACTATCA	427
Db	464	TCGCTCCGCTGGAGAGACTTGTGGAGGACACTGAGACACCAATTCAGACCACTATCA	523
Qy	428	TGGCAAAAATGAGGTTTTCTGGCTCCAAACAGAGAGAGGCGCGCAACCAAGCTCGCC	487
Db	524	TGGCGAAAATGAGGTTTTCTGGCTCCAAACAGAGAGAGGCGCGCAACCAAGCTCGCC	583
Qy	488	TCATCGTATCCAGACCTGGGAGTTGCTGTATGCGAGAGATGGCCCTTTAGCAGCTGG	547
Db	584	TTATCGTATCCAGACCTGGGAGTTGCTGTATGCGAGAGATGGCCCTTTAGCAGCTGG	643
Qy	548	TTTCCACTTTCCTCAGGCGTGATGGGCTCCTCATACGATTTCCAAATCTCTCTTAAGC	607
Db	644	TTTCCACTTTCCTCAGGCGTGATGGGCTCCTCATACGATTTCCAAATCTCTCTTAAGC	703
Qy	608	AGCGGCTCGAGTTCTCTGGTGAATACCTGGAAAGCAAGAAATGCCCTATGGGTTCTCAT	667
Db	704	AGCGGCTCGAGTTCTCTGGTGAATACCTGGAAAGCAAGAAATGCCCTATGGGTTCTCAT	763
Qy	668	ATGACACCCGCTGTTTGTGACTCAACGGTCACTGAGATGACATCCGTTGTGAGAGTCAA	727
Db	764	ATGACACCCGCTGTTTGTGACTCAACGGTCACTGAGAGCGACATCCGTTGTGAGAGTCAA	823
Qy	728	TTTACCAATGTTGTGACTTGGGCGCCGAGCTAGACAGGCCATTAAGGTGCTTCACAGAGC	787
Db	824	TCTACCAATGTTGTGACTTGGGCGCCGAGCTAGACAGGCCATTAAGGTGCTTCACAGAGC	883
Qy	788	GGCTCTATGTGGGGGTCCTGACTTAACCTCCAAAGGCGAGAACTGGCGGTATCGCGGT	847
Db	884	GGCTCTATGTGGGGGTCCTGACTTAACCTCCAAAGGCGAGAACTGGCGGTATCGCGGT	943
Qy	848	GCCGCGGAGCGGGTGTGACCACTAGCTGCGGTAAATACCCCTCACATGCTACTTGAAGG	907
Db	944	GCCGCGGAGCGGGTGTGACCACTAGCTGCGGTAAATACCCCTCACATGCTACTTGAAGG	1003
Qy	908	CCGCTGACGCTGTGAGCTGCCAAGCTCCAGGAGCTGCAAGATGCTGTAAGTGGAGAGC	967
Db	1004	CCTCTGGGCGCTGTGAGCTGCCAAGCTCCAGGAGCTGCAAGATGCTGTAAGTGGAGAGC	1063

Qy	968	ACCTTGTGTTATCTGTGAAGCGCGGGAACCCAGAGACGCGGCAAGCCTACAGTCT	1027
Db	1064	ACCTTGTGTTATCTGTGAGAGCGCGGGAACCCAGAGAGTGGCGGAACCTACAGTCT	1123
Qy	1028	TCACGAGGCTATGACTAGTACTCTGCCCGCCCTGGGACCGCGCCCAACCGGAATACG	1087
Db	1124	TCACGAGGCTATGACTAGTACTCTGCCCGCCCTGGGACCGCGCCCAACCGGAATACG	1183
Qy	1088	ACTTGGAGCTGATAACATCGTGTCTCCAAATGTGCGTGCACAGATGCAATCTGGCA	1147
Db	1184	ACTTGGAGTGTATAACATCATGCTCTCCAAATGTGCGTGCAGCATGCAATCTGGCA	1243
Qy	1148	AAAGGTTGTTACTACCTCACCGTGACCCCAACCGTCCCTTGGCGGGCTGCGTGGGAGA	1207
Db	1244	AAAGGTTGTTACTACCTCACCGTGACCCCAACCGTCCCTTGGCGGGCTGCGTGGGAGA	1303
Qy	1208	CAGTAGGACACTCCAGTCAACCTGCTGGCTAGGCAACATCATCATGTATGCCGCCACTT	1267
Db	1304	CAGTAGGACACTCCCAATCACTCTCTGGCTAGGCAATATCATCATGTATGCCGCCACTT	1363
Qy	1268	TGTGGCAAGGATGATTCTGATGACTCAGTCTCTTCTCCATCTCTTAGCCAGGAGCAAC	1327
Db	1364	TATGGCAAGGATGATTCTGATGACTCATTTCTCTCCATCTCTTAGCCAGGAGCAAC	1423
Qy	1328	TTGAAAAGCCCTGGATTGTGAGATCTACGGGGCTTGTACTCATTGAGCCACTTGACC	1387
Db	1424	TTGAAAAGCCCTGGATTGTGAGATCTACGGGGCTTGTACTCATTGAGCCACTTGACC	1483
Qy	1388	TACCTAGATCATGAAAGCACTCATGCTTACGGGGCTTGTACTCATTGAGCCACTTGACC	1447
Db	1484	TACCTAGATCATGAAAGCACTCATGCTTACGGGGCTTGTACTCATTGAGCCACTTGACC	1543
Qy	1448	CAGTGAGATCAATAGGCTGGCTTCATGCTCAGGAGCTTGGGTTACCACTTGGCGAG	1507
Db	1544	CAGTGAGATCAATAGGCTGGCTTCATGCTCAGGAGCTTGGGTTACCACTTGGCGAG	1603
Qy	1508	TCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTTAAGTTACTGTCCAGGGGGGAGGCGG	1567
Db	1604	TCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTTAAGTTACTGTCCAGGGGGGAGGCGG	1663
Qy	1568	CCATTTGTGGCAAGTACTCTTCAACTGGGCAAGTAAAGCAACAGCTTAAACTCACTCAA	1627
Db	1664	CCATTTGTGGCAAGTACTCTTCAACTGGGCAAGTAAAGCAACAGCTTAAACTCACTCAA	1727
Qy	1628	TTCCGCTGCTGCCGCTGAGACTTGTCCGCTGCTTGGTGGCTTACAGCGGGGAG	168
Db	1724	TCCCGCTGCTGCCGCTGAGACTTGTCCGCTGCTTGGTGGCTTACAGCGGGGAG	178
Qy	1688	ACATATATCACAGCCTGTCTGTCGCCGACCCCG	1721
Db	1784	ACATATATCACAGCCTGTCTGTCGCCGACCCCG	1817

RESULT 4
US-10-492-178-5
; Sequence 5, Application US/10492178
; Publication No. US20040247615A1
; GENERAL INFORMATION:
; APPLICANT: Emini, Emilio A.
; APPLICANT: Kaslow, David C.
; APPLICANT: Bett, Andrew J.
; APPLICANT: Shiver, John W.
; APPLICANT: Nicosia, Alfredo
; APPLICANT: Lahm, Armin
; APPLICANT: Luzzago, Alessandra
; APPLICANT: Cortese, Riccardo
; APPLICANT: Colloca, Stefano
; TITLE OF INVENTION: HEPATITIS C VIRUS VACCINE
; FILE REFERENCE: ITR0015YP
; CURRENT APPLICATION NUMBER: US/10/492,178
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: PCT/US02/32512
; PRIOR FILING DATE: 2002-10-10

; PRIOR APPLICATION NUMBER: 60/363,774
 ; PRIOR FILING DATE: 2002-03-13
 ; PRIOR APPLICATION NUMBER: 60/328,655
 ; PRIOR FILING DATE: 2001-10-11
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 5955

; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: NS cDNA sequence
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(5955)
 US-10-492-178-5

Query Match 88.7%; Score 1543.8; DB 20; Length 5955;
 Best Local Similarity 93.8%; Pred. No. 0;
 Matches 1608; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY	8	GCAATGCTCAATGCTCTATACATGACAGGCGCCCTGATCACACGTCGCTGCGGAGAAA	67
DB	4178	GCTGCTCAATGCTCTACACATGACAGGCGCTTGTATACGCCATGCGCTGGGAGAAA	4237
QY	68	GCAAGCTGCCATCAACGCGCTGAGCAATCTCTGCTGCGTCAACATAACCTGGTCTATT	127
DB	4238	GCAAGCTGCCATCAACGCGCTGAGCAATCTCTGCTGCGTCAACATAACCTGGTCTATT	4297
QY	128	CCACACATCCCGCAGTGCAGCCTGCGGAGAGAGTCACTTTGACAGACTGCAAG	187
DB	4298	CCACACATCTCGCAGCGGAGGCTGCGGAGAGAGTCACTTTGACAGACTGCAAG	4357
QY	188	TCTGTGACGATCAATACCGGAGCTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA	247
DB	4358	TCTGTGACGATCAATACCGGAGCTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA	4417
QY	248	AGGCTAAATGCTCTATCTGTAGAGAGAGCATGCAAGCTGACGCGCCCGCAATTCGGCCAAAT	307
DB	4418	AGGCTAAATGCTCTATCTGTAGAGAGAGCATGCAAGCTGACGCGCCCGCAATTCGGCCAAAT	4477
QY	308	CCAAATTTGGCTATGGGCAAGGAGCGTCCGAGGCTATCCAGCAGGCGGTTAAACCACA	367
DB	4478	CCAAATTTGGCTATGGGCAAGGAGCGTCCGAGGCTATCCAGCAGGCGGTTAAACCACA	4537
QY	368	TCCGCTCCGCTGTGGAAGGACTGCTGGAGGACACTGACACCAATTCAGACCACTCA	427
DB	4538	TCCGCTCCGCTGTGGAAGGACTGCTGGAGGACACTGACACCAATTCAGACCACTCA	4597
QY	428	TGGCAAAAATGAGGTTTTCTCGTCCACAGAGAAAGGAGGCGCGCAACCGCTCGCC	487
DB	4598	TGGCAAAAATGAGGTTTTCTCGTCCACAGAGAAAGGAGGCGCGTAAGCCAGCCCGCC	4657
QY	488	TCATCGTATTTCCAGACCTGGGAGTTGCTGTATGCGAGAAAGATGGCCCTTTACGACGTGG	547
DB	4658	TTATCGTATTTCCAGATCTGGGAGTCCGCTGTATGCGAGAAAGATGGCCCTTTATGATGG	4717
QY	548	TTTCCACTCTTCTCAGGCGGTGATGGCTCTCTATACGGATTCCTCAATCTCTCTAAGC	607
DB	4718	TCTCCACCCCTTCTCAGGCTGATGGCTCTCTATACGGATTCCTCAATCTCTCTAAGC	4777
QY	608	AGCGGCTCAGTCTCTGTTGATACCTGGAAAGCAAGAAATGCGCTATGGCTTCTCAT	667
DB	4778	AGCGGCTCAGTCTCTGTTGATACCTGGAAAGCAAGAAATGCGCTATGGCTTCTCAT	4837
QY	668	ATGACACCGCTGTTTGTACTCAACGCTCACTGAGAAAGTACATCCGTTGTAGAGAGTCAA	727
DB	4838	ATGACACTCGCTGTTTGTACTCAACGCTCACTGAGAAAGTACATCCGTTGTAGAGAGTCAA	4897
QY	728	TTTACCAATGTTGTACTTTGGCCCGCCAGGCTAGACAGGCGCAATAAGGTCGCTCACAGAGC	787
DB	4898	TTTACCAATGTTGTACTTTGGCCCGCCAGGCTAGACAGGCGCAATAAATCGCTCACAGAGC	4957

QY	788	GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGCGGCTATCGCCGT	847
DB	4958	GGCTTTTATATCGGGGTCTCTGACTAAATTCAAAAGGCGAGAACTGCGGTTATCGCCGT	5017
QY	848	GCGCGCGAGCGCGCTGCTGACGACTAGCTGGGTAATACCTCATCTCATCTTGAAGG	907
DB	5018	GCGCGCGAGCGCGCTGCTGACGACTAGCTGGGTAATACCTCATCTCATCTTGAAGG	5077
QY	908	CCGCTGCAGCCTGTCGAGCTGCCAGCTCCAGGACTGCACGATGCTCGTGAATGGAGAG	967
DB	5078	CCTCTGCAGCCTGTCGAGCTGCCAGCTCCAGGACTGCACGATGCTCGTGAACGAGAGG	5137
QY	968	ACCTTGTGCTTATCTGTGAAGCGCGGAAACCCAAAGAGAGACCGCGAAGCCTACGAGTCT	1027
DB	5138	ACCTTGTGCTTATCTGTGAAGCGCGGAAACCCAAAGAGAGACCGCGAAGCCTACGAGTCT	5197
QY	1028	TCACGAGGCTATGACTAGTACTCTGCCCGCTGGGAGCCCGCCCAACCGGANTAGC	1087
DB	5198	TCACGAGGCTATGACTAGTACTCTGCCCGCTGGGAGCCCGCCCAACCGGANTAGC	5257
QY	1088	ACTTCGAGCTGATTAACATCTGTTCTCCTCAATGCTGCGTGCACACGATGATCTGGCA	1147
DB	5258	ACTTCGAGCTGATTAACATCTGTTCTCCTCAATGCTGCGTGCACACGATGATCTGGCA	5317
QY	1148	AAAGGTTGACTTACTCACCCTGACCCCGTCCCTCTTTCGCGGCTGCGTGGGAGA	1207
DB	5318	AAAGGTTGACTTACTCACCCTGACCCCGTCCCTCTTTCGCGGCTGCGTGGGAGA	5377
QY	1208	CAGCTAGGACACATCTCAGTCAACTCTCGCTAGGCAACATCATATGATGCGCCACTT	1267
DB	5378	CAGCTAGGACACATCTCAGTCAACTCTCGCTAGGCAACATCATATGATGCGCCACTT	5437
QY	1268	TGTGGGCAAGGATGATCTGATGACTCTTCTTCTCCATCTTTCAGCCAGGAGCAAC	1327
DB	5438	TGTGGGCAAGGATGATCTGATGACTCTTCTTCTCCATCTTTCAGCCAGGAGCAAC	5497
QY	1328	TTGAAAAAGCCCTGGATTGTGAGTCTACGGGCTTGTACTTCCATTGAGCCACTTGACC	1387
DB	5498	TTGAAAAAGCCCTGGATTGTGAGTCTACGGGCTTGTACTTCCATTGAGCCACTTGACC	5557
QY	1388	TACTCAGATCATTTGAACGACTCCATGGTCTTAGCGCAATTTTCACTCCATAGTTACTCTC	1447
DB	5558	TACTCAGATCATTTGAACGACTCCATGGTCTTAGCGCAATTTTCACTCCATAGTTACTCTC	5617
QY	1448	CAGTGAATCAATAGGTTGCTTCAATGCTCAGGAGCTTGGGTACACCCCTTGGAG	1507
DB	5618	CAGTGAATCAATAGGTTGCTTCAATGCTCAGGAGCTTGGGTACACCCCTTGGAG	5677
QY	1508	TCTGAGACATCGGGCCAGAGTGTCCGCGTAACTTCTCCAGGGGGGAGGCGG	1567
DB	5678	TCTGAGACATCGGGCCAGAGTGTCCGCGTAACTTCTCCAGGGGGGAGGCGG	5737
QY	1568	CCATTGTGGCAAGTACCTCTTCAACTGGGCAAGTAAGGACCAAGCTTAAACTCACTCCAA	1627
DB	5738	CCATTGTGGCAAGTACCTCTTCAACTGGGCAAGTAAGGACCAAGCTTAAACTCACTCCAA	5797
QY	1628	TTCCGCTGCTGCTCCGCTGAGTGTTCGCGCTGGTTCGTTGCTGGCTACAGCGGGAG	1687
DB	5798	TTCCGCTGCTGCTCCGCTGAGTGTTCGCGCTGGTTCGTTGCTGGCTACAGCGGGAG	5857
QY	1688	ACATATATCAGAGCTGCTCGTCCGCGACCCCGC	1722
DB	5858	ACATATATCAGAGCTGCTCGTCCGCGACCCCGC	5892

RESULT 5
 US-10-492-178-2
 ; Sequence 2, Application US/10492178
 ; Publication No. US20040247615A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Emini, Emilio A.
 ; APPLICANT: Kaslow, David C.
 ; APPLICANT: Bett, Andrew J.

; APPLICANT: Shiver, John W.
; APPLICANT: Nicosia, Alfredo
; APPLICANT: Lahm, Armin
; APPLICANT: Luzzago, Alessandra
; APPLICANT: Cortese, Riccardo
; APPLICANT: Colloca, Stefano
; TITLE OF INVENTION: HEPATITIS C VIRUS VACCINE
; FILE REFERENCE: ITR0015YP
; CURRENT APPLICATION NUMBER: US/10/492,178
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: PCT/US02/32512
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/363,774
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/328,655
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 5965
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Non-optimized cDNA sequence encoding SEQ. ID. NO.
; OTHER INFORMATION: 1

US-10-492-178-2

Query Match	88.4%;	Score 1537.4;	DB 20;	Length 5965;
Best Local Similarity	93.5%;	Pred. No. 0;		
Matches 1604;	Conservative 0;	Mismatches 111;	Indels 0;	Gaps 0;
QY	8	GCATGTCAATGTCCTATACATGGACAGCGCGCTCATCACACCGTGCCTCGGAGGAAA	67	
DB	4184	GCTGTCAATGTCCTACACATGGACAGCGCGCTTGATCACGCCATGCGCTCGGAGGAAA	4243	
QY	68	GCAAGCTCCCATCAACGGCTGAGCAACTCTTCTCGTGGGTCACCATAACTGCTCTATT	127	
DB	4244	GCAAGCTGCCATCAACGGCTTGAGCAACTCTTCTCGTGGGTCACCATAACTGCTCTATT	4303	
QY	128	CCACAACATCCCGCAGTCAAGCCTCGGCGAGAGAGAGGTCACTTTTGACAGACTGCAAG	187	
DB	4304	CCACAACATCTCGAGCGCAGCGCTCGGCGAGAGAGAGGTCACTTTTGACAGACTGCAAG	4363	
QY	188	TCCTGGAGCATCATTAACCGGACGTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTGA	247	
DB	4364	TCCTGGAGCACCATTACCGGACGTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTGA	4423	
QY	248	AGGCTAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCCCGCGCATTTGCGCCAAAT	307	
DB	4424	AGGCTAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCCCGCGCATTTGCGCCAAAT	4483	
QY	308	CCAAATTTGGCTATGGGCGAAGGACGTCCGAGCCTATCCAGCAGGCGCGTTAACCAACA	367	
DB	4484	CCAAGTTTGGCTATGGGCGAAGGACGTCCGAACTATCCAGCAGGCGCGTTAACCAACA	4543	
QY	368	TCCGCTCCGCTGGAGAGACTTGTGGAGACACTGACACACCAATTCAGACCAACCATCA	427	
DB	4544	TCCATCTCCGCTGGAGAGACTTGTGGAGACACTGTCACCAATTCAGACCAACCATCA	4603	
QY	428	TGGCAAAAAATGAGGTTTCTGCGTCCAAACAGAGAAAGGAGCGCGCAAAACAGCTCGCC	487	
DB	4604	TGGCAAAAAATGAGGTTTCTGCTGTCACCAACAGAGAAAGGAGCGCGTAAAGCGCCGCC	4663	
QY	488	TCATCGTATTTCCAGACTGGGAGTTCGTGTATCGAGAGAGATGGCCCTTTACGAGCTGG	547	
DB	4664	TTATCGTATTTCCAGACTGGGAGTTCGTGTATCGAGAGAGATGGCCCTTTATGATGTGG	4723	
QY	548	TTTCCACTCTTCTCAGGCGCTGATGGGCTCCTCATACAGATTCCAATACTCTCTAAAGC	607	
DB	4724	TTTCCACCCTTCTCTAGGTCGTGATGGGCTCTCATACAGATTCCAATACTCTCTCTGGGC	4783	
QY	608	AGCGGCTCGAGTTCTCTGTAATACCTGGAAAGCAAGAAATGSCCTATGCGGCTTCTCAT	667	

4784	AGCGAGTCGAGTTCTCTGTTGAATACCTGGAAATCAAGAAAAACCCCATGGGCTTTTCAT	4843	DB
668	ATGACACCCGCTGTTTGTGACTCAACCGGTCACTGAGAAATGACATCCGTGTGAGGAGTCAA	727	QY
4844	ATGACACTCGCTGTTTCGACTCAACCGGTCAACCGGTCAACCGGTCAACCGGTCAAC	4903	DB
728	TTTACCAATGTTGTGACTTTGGGCCCCGAAAGCTAGACAGAGCCATAGGTCCTCACAAGAC	787	QY
4904	TTTACCAATGTTGTGACTTTGGGCCCCGAAAGCCAGACAGAGCCATATAAATCGCTCACA	4963	DB
788	GGCTCTATGTCGGGGTCCATGACTAACTCCAAAGGGCAGAACTGCGGCTATTCGCCGT	847	QY
4964	GGCTTTATATCGGGGTCTCTGACTAATTTCAAAGGGCAGAACTCGCGGTATTCGCCGT	5023	DB
848	GCCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAAATACCTCTACATGCTACTTGAAGG	907	QY
5024	GCCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAAACCTCTACATGCTACTTGAAGG	5083	DB
908	CCGCTGCAGCTCTCGAGCTGCCAAGCTCCAGACTGCACGATGCTCGTGAATGGAGACG	967	QY
5084	CCCTGCGAGCTCTCGAGCTGCCAAGCTCCAGACTGCACGATGCTCGTGAACGCGCGCG	5143	DB
968	ACCTTGTCGTTATCTGTGAAAGCGCGGAAACCCAAAGAGGACGCGGCAAGCTTACGAGTCT	1027	QY
5144	GCCTTGTCGTTATCTGTGAAAGCGCGGAAACCCAAAGAGGACGCGGCGGAGCTTACGAGTCT	5203	DB
1028	TCACGAGGCTATGACTAGGTACTCTGCCCGCTCGGGGACCGCGCCCAACCGGAATACG	1087	QY
5204	TCACGAGGCTATGACTAGGTACTCTGCCCGCTCGGGGACCGCGCCCAACCGGAATACG	5363	DB
1088	ACTTGGAGCTGATAACAATCGTGTTCCTCAATGTGTCGTCGACACGATGCTATCGGCA	1147	QY
5264	ACTTGGAGCTGATAACAATCATGTTCCTCAATGTGTCGTCGACGATGCTATCGGCA	5323	DB
1148	AAAGGTTGTTACTACCTCACCGGTGACCCCAACCGTCCCGCTTCGCGCGGCTGCTGGAGA	1207	QY
5324	AAAGGTTGTTACTACCTCACCGGTGATCCCAACCGTCCCGCTTCGCGCGGCTGCTGGAAA	5383	DB
1208	CAGTAGGCACTCTCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGCGCCACTT	1267	QY
5384	CAGTAGGCACTCTCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGCGCCACTT	5443	DB
1268	TGTCGGCAAGGATGATTTGATGACTCACTTCTTCTCCATCTTAGCCACGAGCAAC	1327	QY
5444	TGTCGGCAAGGATGATTTGATGACTCACTTCTTCTCCATCTTAGCCACGAGAGCAAC	5503	DB
1328	TTGAAAAAGCCCTGGATTGTGAGATCTACCGGGGCTTGTACTCATTGAGCCACTTGACC	1387	QY
5504	TTGAAAAAGCCCTGGACTGCGAGATCTACCGGGGCTTGTACTCATTGAGCCACTTGACC	5563	DB
1388	TACCTCAGATCATTTGAAACGACTCCATGGTCTTAGCGCATTTTCACTCATAGTTACTCTC	1447	QY
5564	TACCTCAGATCATTTGAAACGACTCCATGGCTTAGCGCATTTTCACTCATAGTTACTCTC	5623	DB
1448	CAGGTGAGATCAATAGGCTGCTTTCATGCTCAGGAAGCTTGGGTTACCACTTGGGAG	1507	QY
5624	CAGGTGAGATCAATAGGCTGCTTTCATGCTCAGGAAGCTTGGGTTACCACTTGGGAG	5683	DB
1508	TCTCGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGGCGG	1567	QY
5684	TCTCGAGACATCGGGCCAGGAGCGTCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCGG	5743	DB
1568	CCATTTGTGGCAAGTACTCTTTCAACTGGGCGAGTAAAGCAAGCTTTAAACTCACTCAA	1627	QY
5744	CCACTTGTGGCAAGTACTCTTTCAACTGGGCGAGTAAAGCAAGCTTTCAACTCACTCAA	5803	DB
1628	TTCCGGCTGCGCTCCCGCTGGAATTTGTCCGGCTGGTGTGTTGCTGGCTACAGCGGGGAG	1687	QY
5804	TTCCGGCTGCGCTCCCGCTGGAATTTGTCCGGCTGGTGTGTTGCTGGCTACAGCGGGGAG	5863	DB
1688	ACATATATCACAGCTGTCTCGTCCCGGACCCCGC	1722	QY
5864	ACATATATCACAGCTGTCTCGTCCCGGACCCCGC	5898	DB

Db 7001 CCACTTGTGGCAAGTACCTCTTCAACTGGCGAGTGAAGACCAAACTCAACTCACTCCAA 7060
Qy 1628 TTCGGGCTGGTCCCGCTGGACCTTGTCCGGCTGGTTCGTTCGTGGCTACAGCGGGGAG 1687
Db 7061 TCCGGGCTGGCTCCAGCTGGACTTGTCCGGCTGGTTCGTTCGTTCAGCGGGGAG 7120
Qy 1688 ACATATATCACAGCTGTCTCGTGCCCGACCCCGC 1722
Db 7121 ACATATATCACAGCTGTCTCGTGCCCGACCCCGC 7155

RESULT 7
US-10-259-275-41
; Sequence 41, Application US/10259275
; Publication No. US20030125541A1
; GENERAL INFORMATION:
; APPLICANT: Lemon, Stanley M.
; APPLICANT: Yi, Minkyung
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007 0120
; CURRENT APPLICATION NUMBER: US/10/259,275
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US 60/171,909
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/747,419
; PRIOR FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/325,236
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/338,123
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 41
; LENGTH: 6189
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding hepatitis C virus polyprotein
; OTHER INFORMATION: derived from Con1
US-10-259-275-41

Query Match 87.8%; Score 1527.8; DB 15; Length 6189;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1598; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Qy 8 GCATGTCATATGTCCTATACATGACAGCGGCGCTGATCACACCGTGGCTGGAGGAAA 67
Db 4178 GCTGCTCGATGCTTACATACATGACAGCGGCGCTGATCACGCGATGCGCTGGAGGAAA 4237
Qy 68 GCAAGTGGCCATCAACGGCTGAGCAACTCCTTGTGCTGCGTCACCATAACCTGGTCTATT 127
Db 4238 CCAAGTGGCCATCAATGACTGAGCAACTCTTGTCTCGTCAACCACAACTTGGTCTATG 4297
Qy 128 CCACAACTCCCGAGTGAAGCTGGCGGAGAGAGAGTCACTTTGACAGACTGCGAG 187
Db 4298 CTACAACTCTCGACGCGAAGCTTGGCGGAGAGAGTCACTTTGACAGACTGCGAG 4357
Qy 188 TCTGAGACATCATTTACCGGAGCTGCTCAAGGAGATGAAGCGAGCGGCTCCACAGTGA 247
Db 4358 TCCTGACGACCACTACCGGAGCTGCTCAAGGAGATGAAGCGAGCGGCTCCACAGTTA 4417
Qy 248 AGCTAACTGCTATCTGTAGAAAGCATGCAAGCTGACGCGCCCGCATTCGGCCAAAT 307
Db 4418 AGCTAACTTCTATCCGTGGAGAGAGCTGTAGCTGAGCGCCCGCCACATTCGGCCAGAT 4477
Qy 308 CCAAAATTTGGCTATGGGCAAGAGAGCTCGGAGCGCTATCCAGAGCGGCGGTAAACCA 367
Db 4478 CTAAATTTGGCTATGGGCAAGAGAGCTCCGGAACCTATCCAGAGCGGCGGTAAACCA 4537
Qy 368 TCCGCTCCGTGGAGAGACTTCTGTGAGGAGACTGACAGACACTGACACACCAATTCAGAC 427
Db 4538 TCCGCTCCGTGGAGAGACTTCTGTGAGAGAGACTGAGACACCAATTCAGACACCAATCA 4597

Qy 428 TGGCAAAAATCAGGTTTCTGCGTCCACACAGAGAAAGAGGCGCGCAAAACAGCTCGCC 487
Db 4598 TGGCAAAAATCAGGTTTCTGCGTCCAAACAGAGAGAGGGGGCGCGCAAGCCAGCTCGCC 4657
Qy 488 TCATCGTATTTCCAGACCTGGGAGTTGGTGTATGCGAGAAATGCGCCCTTTACGAGTGG 547
Db 4658 TTATCGTATTTCCAGATTTGGGGGTTGGTGTGCGAGAAATGCGCCCTTTACGATGTGG 4717
Qy 548 TTTTCCACTCTTCTCAGGCGGTGATGGGCTCTCATACGGAATTCATAATCTCTCTTAAGC 607
Db 4718 TCTCCACCTCTCCCTCAGGCGGTGATGGGCTCTTCATACGGAATTCATAATCTCTCTGGAC 4777
Qy 608 AGCGGTGCGAGTTCTCTGCTGGAATACCTGGAAGCAAGAAATGCGCTATGGGCTTCTCAT 667
Db 4778 AGCGGTGCGAGTTCTCTGCTGGAATGCGCTGGAAGCAAGAAATGCGCTATGGGCTTCTCGCAT 4837
Qy 668 ATGACACCGGCTGTTTGTACTCAACGGTCACTGAGAAATGACATTCGGTGTGGAGGACTCAA 727
Db 4838 ATGACACCGGCTGTTTGTACTCAACGGTCACTGAGAAATGACATTCGGTGTGGAGGACTCAA 4897
Qy 728 TTTTACCAATGTTGTGACTTTGGCGCCCGAAGCTAGACAGGCCATAAGGTCGCTCACAGAGC 787
Db 4898 TCTACCAATGTTGTGACTTTGGCGCCCGGAGCCAGACAGGCCATAAGGTCGCTCACAGAGC 4957
Qy 788 GGCTCTATGTCGGGGTCCCATGACTTAACTTCCAAAGGGCAGAAATCTCGGGTATTCGGCGGT 847
Db 4958 GGCTTTTACATCGGGGGCGCCCTGACTAAATCTTAAAGGGCAGAAATCTCGGGTATTCGGCGGT 5017
Qy 848 GCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAAATACCTTCACATGCTACTCTTGAAGG 907
Db 5018 GCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAAATACCTTCACATGCTACTCTTGAAGG 5077
Qy 908 CCGCTCAGCGCTCTCGAGCTGCCAAGCTCCAGGACTGACAGATGCTCGTGAATGGAGAGC 967
Db 5078 CCGCTCAGCGCTCTCGAGCTGCCAAGCTCCAGGACTGACAGATGCTCGTGAATGGAGAGC 5137
Qy 968 ACCTTGTGCTTATCTGTGAAGCGCGGAAACCCAAAGAGGACGCGGAAGCTTACGAGTCT 1027
Db 5138 ACCTTGTGCTTATCTGTGAAGCGCGGAAACCCAAAGAGGACGAGGCGGAGCTTACGCGCT 5197
Qy 1028 TCACGAGGCTATGACTAGTACTCTGCGCCCGCTGGGAGACCGCGCCCAACCGGAATACG 1087
Db 5198 TCACGAGGCTATGACTAGTACTCTGCGCCCGCTGGGAGACCGCGCCCAACCCAGGAATACG 5257
Qy 1088 ACTTGGAGCTGATTAACATCGTGTTCCTCAATGTCGTCGTCGACACGATGCTATCTGGCA 1147
Db 5258 ACTTGGAGTTGATAAACATCATGCTCTCCAAATGTCAGTTCGCGACGATGCTATCTGGCA 5317
Qy 1148 AAGGGTGTACTACTCTCACCGGTGACCCCAACCGTCCCGCTTGGCGGAGCTTGGTGGAGGA 1207
Db 5318 AAGGGTGTACTATCTCACCGGTGACCCCAACCGCTTGGCGGAGCTTGGTGGAGGA 5377
Qy 1208 CAGCTAGGCGACACTCCAGTCAAACTCTCGGCTAGGCAACATCATGATGTCGCGCCACTT 1267
Db 5378 CAGCTAGGCGACACTCCAGTCAAACTCTCGGCTAGGCAACATCATGATGTCGCGCCACTT 5437
Qy 1268 TGTGGCAAGGATGATTTCTGATGACTCATCTTCTCTCATGCTTCTAGCCAGGAGCAAC 1327
Db 5438 TGTGGCAAGGATGATTTCTGATGACTCATTTCTCTCATGCTTCTAGCCAGGAGCAAC 5497
Qy 1328 TTGAAGAAGCCCTGGATGTCAGATCTACGGGGCTTGTACTTCATTTAGGCACTTTCAGC 1387
Db 5498 TTGAAGAAGCCCTGGATGTCAGATCTACGGGGCTTGTACTTCATTTAGGCACTTTCAGC 5557
Qy 1388 TACCTCAGATCATTTGAAGAGCTCCATGGTCTTACGGCATTTTCACTCCATATGAGTACTTCT 1447
Db 5558 TACCTCAGATCATTTGAAGAGCTCCATGGCTTACGGCATTTTCACTCCATATGAGTACTTCT 5617
Qy 1448 CAGGTGAGATCAATAGGGTGGCTTCAATGCTCAGAAAGCTTGGGGTACCACTTTCGAG 150
Db 5618 CAGGTGAGATCAATAGGGTGGCTTCAATGCTCAGAAAGCTTGGGGTACCACTTTCGAG 567
Qy 1508 TCTGGAGACATGGGCCAGAAAGTGTCCGGCTTAAGTTACTTCTCCAGGGGGGGCGG 156


```
Db 5678 TCTGGAGACATCGGGCCAGAGTGTCGGGCTAGGCTACTGTCACGAGGGGGAGGGCTG 5737
Oy 1568 CCAATTGTGGAGTACCTCTTCAACTGGGAGTAAAGGACCAAGCTTAAACTCACTCCAA 1627
Db 5738 CCACTTGTGGCAAGTACCTCTTCAACTGGGAGTAAAGGACCAAGCTTAAACTCACTCCAA 5797
Oy 1628 TTCCGGCTCGCTCCGGCTGGACCTTGTCCGGCTGGTTCCTTCTGCTGGCTACAGCGGGGAG 1687
Db 5798 TCCGGCTCGCTCCGGCTGGATTTATCCAGCTGGTTCGTTGCTGTTACAGCGGGGAG 5857
Oy 1688 ACATATATCACAGCTGTCTCGTGGCCGACCCCGC 1722
Db 5858 ACATATATCACAGCTGTCTCGTGGCCGACCCCGC 5892

RESULT 8
US-11-006-313-41
; Sequence 41, Application US/11006313
; Publication No. US20050153281A1
; GENERAL INFORMATION:
; APPLICANT: Lemon, Stanley M.
; APPLICANT: Yi, Minkyung
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007 G121
; CURRENT APPLICATION NUMBER: US/11/006,313
; PRIOR FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US 60/171,909
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 10/259,275
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US 09/747,419
; PRIOR FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/325,236
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/338,123
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41
; LENGTH: 6189
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding hepatitis C virus polypeptide
; OTHER INFORMATION: derived from Con1
US-11-006-313-41

Query Match 87.8%; Score 1527.8; DB 24; Length 6189;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1598; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Oy 8 GCATGTCAATGTCTTATACATGACGAGGGCCCTGATCACACCGTGCCTCGGAGGAAA 67
Db 4178 GCTGCTCGATGTCTTACATGACGAGGGCCCTGATCACGCCATGCGCTGGGAGGAAA 4237
Oy 68 GCAAGCTGCCATCAAGCGGTGAGCAACTCTTCTGCTGGTCAACATTAACCTGFTATT 127
Db 4238 CCAAGCTGCCATCAATGACCTGAGCAACTCTTCTGCTGGTCAACCAACTTGGTCTATG 4297
Oy 128 CCACAACTCCGGAGTCAAGCTCGGCGAGAGAGAGTCACTTTGACAGACTGCAAG 187
Db 4298 CTCACAACTCTGGAGCGAGCTCGGCGAGAGAGAGTCACTTTGACAGACTGCAAG 4357
Oy 188 TCCTGGAGATCATTAACCGGAGCTGCTCAAGGAGATCAAGCGAGAGCGTCCACAGTGA 247
Db 4358 TCCTGGAGAGCACTACCGGAGCTGCTCAAGGAGATCAAGCGAGAGCGTCCACAGTGA 4417
Oy 248 AGGCTAAACTGCTATCTGTAGAGAGCATGCAAGCTGACGCCCGCCGATTCGGGCCAAAT 307
Db 4418 AGGCTAAACTTCTATCCGTGGAGGAGGCTGTAGAGTGAACCCCGCCAGATTCGGCCAGAT 4477
Oy 308 CCAATTGTGGCTATGGGGCAAGGAGCTGCGGAGGCTATCCAGCAGGGGCGTTAAACCA 367
```

```
Db 4478 CTAATTTGGCTATGGGGCAAGGAGCTCCGGAACCTATCCAGCAAGGCGGTTAAACCA 4537
Oy 368 TCCGCTCCGCTGGAAAGGACTTGTGGAGGACACTGACACACCAATTCAGACACCATCA 427
Db 4538 TCCGCTCCGCTGGAAAGGACTTGTGGAGGACACTGAGACACCAATTCAGACACCATCA 4597
Oy 428 TGGCAAAATAGAGTTTCTGCTGCCAAACAGAGAAAGGAGCGGCAAAACAGCTCGCC 487
Db 4598 TGGCAAAATAGAGTTTCTGCTGCCAAACAGAGAAAGGAGCGGCGCAAGCAGCTCGCC 4657
Oy 488 TCATCGTATTCAGACACTCGGAGTTCGTGTATGGGAGAAATGCGCCCTTTACGAGTGG 547
Db 4658 TTATCGTATTCAGAGTTTGGGGGTTCTGTGTGCGAGAAATGCGCCCTTTACGATGG 4717
Oy 548 TTTCCACTCTTCTCAGGCGGTGATGGCTCTCATAGGATTCCTAATCTCTCCTTAAGC 607
Db 4718 TCTCACCTCTCAGGCGGTGATGGCTCTTATACGGATTCCTAATCTCTCCTTAAGC 4777
Oy 608 AGCGGTCGAGTTCCTGCTGAATACCTGGAAAGCAAAAGAAATGCGCTATGGGCTTCTCAT 667
Db 4778 AGCGGTCGAGTTCCTGCTGAATACCTGGAAAGCAAAAGAAATGCGCTATGGGCTTCTCAT 4837
Oy 668 ATGACACCGCTGTTTGAATCACTCAACGGTCACTGAGAAATGACATCCGTGTTGAGAGTCA 727
Db 4838 ATGACACCGCTGTTTGAATCACTCAACGGTCACTGAGAAATGACATCCGTGTTGAGAGTCA 4897
Oy 728 TTTACCAATGTTGACTTGGCCCCCGAGTACAGAGGCCATAGGTGCTCTACAGAGC 787
Db 4898 TCTACCAATGTTGACTTGGCCCCCGAGTACAGAGGCCATAGGTGCTCTACAGAGC 4957
Oy 788 GGCTCTATGTCGGGGGTCCTAGTACTAGCTCGGTAATACCTCACAATGCTACTTTGAAG 907
Db 4958 GGCTTTATCATCGGGGCCCCCTGACTTAATTTCTAAGGGGAGAACTGGGGCTATCGCGGT 5017
Oy 848 GCCGCGGAGCGGCTGCTGACGACTAGCTCGGTAATACCTCACAATGCTACTTTGAAG 907
Db 5018 GCCGCGGAGCGGCTGCTGACGACTAGCTCGGTAATACCTCACAATGCTACTTTGAAG 5077
Oy 908 CCGCTGAGGCTGTCGAGCTGCCAAGCTCCAGGACTGACGATGCTCGTGAATGGAGAGC 967
Db 5078 CCGCTGCGGCTGTCGAGCTGCGAAGCTCCAGGACTGACGATGCTCGTGAATGGAGAGC 5137
Oy 968 ACCTTGTGTTATCTGTGAAGCGGGGAAACCCAGAGGAGCGGCGAAGCTAGAGTCT 1027
Db 5138 ACCTTGTGTTATCTGTGAAGCGGGGAAACCCAGAGGAGCGGCGAAGCTAGAGTCT 5197
Oy 1028 TCAGGAGGCTATGACTAGGTACTCTGCCCGCTCGGGAGCCCGCCCAACCGCAATACG 1087
Db 5198 TCAGGAGGCTATGACTAGGTACTCTGCCCGCTCGGGAGCCCGCCCAACCGCAATACG 5257
Oy 1088 ACTTGAGCTGATTAACATCGTGTTCCTCCAATGTGTCGCTCGCACAGATGCTCTGGCA 1147
Db 5258 ACTTGAGCTGATTAACATCATGCTCTCTCCAATGTGTCGCTCGCACAGATGCTCTGGCA 5317
Oy 1148 AAAGGGTGTACTACTCTCAACCGTGAACCCCAACCGTCCCGCTTGGCGGGGCTCGGTGGAGA 1207
Db 5318 AAAGGGTGTACTACTCTCAACCGTGAACCCCAACCGTCCCGCTTGGCGGGGCTCGGTGGAGA 5377
Oy 1208 CAGTAGGACACTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGGCGCCACTT 1267
Db 5378 CAGTAGGACACTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGGCGCCACTT 5437
Oy 1268 TGTGGGCAAGGATGTTCTGATGACTCACTTCTTCTCATCTCTTCTAGCCAGGAGCAAC 1327
Db 5438 TGTGGGCAAGGATGTTCTGATGACTCACTTCTTCTCATCTCTTCTAGCTCAGGAAAC 5497
Oy 1328 TTGAAAAGCCCTGGATTGTGAGATCTACCGGGGCTTGTACTCATTGAGCCACTTGACC 1387
Db 5498 TTGAAAAGCCCTGGATTGTGAGATCTACCGGGGCTTGTACTCATTGAGCCACTTGACC 5557
Oy 1388 TACCTCAGATCATGAAAGCACTCATGCTCTTGTAGGCAATTTTCTCATCTCATGTACTCTC 1447
```


Db	5558	TACCTCAGATCAATTCACAGACTCCATGCGCCTTTAGCGCATTTTTCACACTCCATATGTTACTCTC	5617
Qy	1448	CAGGTGAGATCAATAGGTTGGCTTCATGCTCCAGAAAGCTTTGGGGTACCAACCTTTGCGAG	1507
Db	5618	CAGGTGAGATCAATAGGTTGGCTTCATGCTCAGGAAACTTTGGGGTACCGCCCTTGGAG	5677
Qy	1508	TCTGGAGACATCGGGCCAGAAAGTTCGCGCTTAAGTTACTGTCCAGGGGGGAGGCGCG	1567
Db	5678	TCTGGAGACATCGGGCCAGAAAGTTCGCGCTTAGGCTACTGTCCAGGGGGGAGGCGCTG	5737
Qy	1568	CCATTGTGGCAGTACCTTCCTCACTGGGCAAGTAAAGCAAGCTTAACCTCACTCCAA	1627
Db	5738	CCACTTGTGGCAAGTACCTTCCTCACTGGGCAAGTAAAGCAAGCTTAACCTCACTCCAA	5797
Qy	1628	TTCCGGCTCGTCCCGCTGGACTTGTTCGGCTGCTTGTGCTGGCTTACAGCGGGGAG	1687
Db	5798	TCCGGCTCGTCCCGCTGGACTTGTTCGGCTGCTTGTGCTGGCTTACAGCGGGGAG	5857
Qy	1688	ACATATATCACAGCTGTCTCGTGGCCGACCCCGC	1722
Db	5858	ACATATATCACAGCTGTCTCGTGGCCGACCCCGC	5892
RESULT 9			
US-10-434-842-16			
; Sequence 16, Application US/10434842			
; Publication No. US2004000549A1			
; GENERAL INFORMATION:			
; APPLICANT: Bichko, Vadim			
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REP			
; FILE REFERENCE: 0342/1H395US3			
; CURRENT APPLICATION NUMBER: US/10/434,842			
; CURRENT FILING DATE: 2003-05-09			
; PRIOR APPLICATION NUMBER: US 10/233,307			
; PRIOR FILING DATE: 2002-08-28			
; PRIOR APPLICATION NUMBER: US 10/005,469			
; PRIOR FILING DATE: 2001-11-07			
; PRIOR APPLICATION NUMBER: US 60/245,866			
; PRIOR FILING DATE: 2000-11-07			
; NUMBER OF SEQ ID NOS: 17			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 16			
; LENGTH: 7989			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: FCA4 Replicon Sequence			
US-10-434-842-16			
Query Match 87.8%; Score 1527.8; DB 17; Length 7989;			
Best Local Similarity 93.2%; Pred. No. 0;			
Matches 1598; Conservative 0; Mismatches 117; Indels 0; Gaps 0;			
Qy	8	GCATGTCAATGTCTATACATGACAGAGGGCCCTGATCACCCGTGCGCTGGGAGGAAA	67
Db	5975	GCTGCTCGATGTCTACATACATGACAGAGGGCCCTGATCACCCATGCGCTGCGGAGGAAA	6034
Qy	68	GCAAGCTGCCCATCAACCGCTCAGCAACTCTCTTGTGGTGCACCACTAACCTGGTCTATT	127
Db	6035	CCAAAGCTGCCCATCAATGACTGAGCAACTCTTGTCTCGTCCACCAACTTGGTCTATG	6094
Qy	128	CCACAAATCCCGCAGTGCAGCTCGCGCAGAGAAAGGTCACTTTGACAGACTGCAAG	187
Db	6095	CTACAAATCTCGCAGCGCAAGCTCGCGCAGAGAAAGGTCACTTTGACAGACTGCAAG	6154
Qy	188	TCCTGGACGATCATACCGGAGCTGCTCAAGAGATGAAGGGAAGGCTCCACAGTGA	247
Db	6155	TCCTGGACGACCACTACCGGAGCTGCTCAAGAGATGAAGGGAAGGCTCCACAGTGA	6214
Qy	248	AGCTTAAATGCTTCTCTCAACAAAGCATGCAAGCTCACTCCCTCTCGCATTCGGCCAAAT	307

```
Db 7355 TACCTCAGATCAATTCACGACTCCATGGCTTAGCCATTTTACCTCCATAGTACTCTC 7414
Qy 1448 CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAAGTCTGGGGTACACCTTTGCGAG 1507
Db 7415 CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAATCTTGGGGTACCGCTTGGCAG 7474
Qy 1508 TCTGGAGACATCGGGCCAGAACTGTCGGCGCTAAGTTACTGTCCAGGGGGGAGGGCGG 1567
Db 7475 TCTGGAGACATCGGGCCAGAACTGTCGGCGCTAAGTTACTGTCCAGGGGGGAGGGCTG 7534
Qy 1568 CCAATTTGTGGCAAGTACCTCTTCAACTGGGCAAGTAAAGTAAAGTCACTCCAA 1627
Db 7535 CCACTTTGTGGCAAGTACCTCTTCAACTGGGCAAGTAAAGTAAAGTCACTCCAA 7594
Qy 1628 TTCGGGCTCGCTCCCGCTGGAGCTTGTCCGGCTGGTTCGTTGCTACAGCGGGGAG 1687
Db 7595 TCCCGCTCGCTCCCGCTGGAGCTTGTCCGGCTGGTTCGTTGCTACAGCGGGGAG 7654
Qy 1688 ACATATATCACAGCTGCTGTCGGCCGACCCCGC 1722
Db 7655 ACATATATCACAGCTGCTGTCGGCCGACCCCGC 7689

RESULT 10
US-10-639-150-1
; Sequence 1, Application US/10639150
; Publication No. US20040121975A1
; GENERAL INFORMATION:
; APPLICANT: BRISTOL-MYERS SQUIBB COMPANY
; TITLE OF INVENTION: HEPATITIS C VIRUS ASSAYS
; FILE REFERENCE: D0224 NP
; CURRENT APPLICATION NUMBER: US/10/639,150
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: US 60/402,661
; PRIOR FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 7989
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: HCV Replicon
US-10-639-150-1

Query Match 87.8%; Score 1527.8; DB 19; Length 7989;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1598; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Qy 8 GCATGTCATATGTCCTATACATGGACAGGGCCCTGATCACACCGTGGCGCTGGGAGGAAA 67
Db 5978 GTGCTCGATGTCCTTACACATGGACAGGGCCCTGATCACGCGCATGGCGCTGGGAGGAAA 6037

Qy 68 CAAAGTGGCCATCAACGGCTGAGCAACTCTCTGCTCGTCGTCACCATAACTGGTCTATT 127
Db 6038 CCAAGTGGCCATCAATGACTGAGCAACTCTTCTGCTCGTCACCAAACTTGGTCTATG 6097

Qy 128 CCACAAATCCGCGAGTCAAGCTTGGCGGAGAAAGGTCACTTTTGAAGACTTGAAG 187
Db 6098 CTCACAAATCTCGCAGCGCAAGCTTGGCGGAGAAAGGTCACTTTTGAAGACTTGAAG 6157

Qy 198 TCCTGGAGCATCATTTACGGGAGCTGCTCAAGGAGATGAAGCGGAAAGCGTCCACAGTGA 247
Db 6158 TCCTGGAGCATCATTTACGGGAGCTGCTCAAGGAGATGAAGCGGAAAGCGTCCACAGTGA 6217

Qy 248 AGGCTAAACTGCTATCTTAGAAGACATGCAAGCTGACGCGCCCGCATTTGGGCCAAAT 307
Db 6218 AGGCTAAACTGCTATCTTAGAAGACATGCAAGCTGACGCGCCCGCATTTGGGCCAAAT 6277

Qy 308 CCAAAATTTGGCTATTTGGGCAAGGAGCTGCGGAGCTTATCCAGCAGGCGCGTGAACACA 367
Db 6278 CTAATTTGGCTATTTGGGCAAGGAGCTGCGGAGCTTATCCAGCAGGCGCGTGAACACA 6337
```

```
Qy 368 TCCGCTCGCTGTGGAGGACCTTGTGTGGAGGACACTGACACACCAATTCAGACCACCATCA 427
Db 6338 TCCGCTCGCTGTGGAGGACCTTGTGTGGAGGACACTGACACACCAATTCAGACCACCATCA 6397
Qy 428 TGGCAAAAATGAGGTTTTTCTGCTCCAAACAGAGAAAGGAGCGCGCAACACGAGCTCGC 487
Db 6398 TGGCAAAAATGAGGTTTTTCTGCTCCAAACAGAGAAAGGAGCGCGCAACACGAGCTCGC 6457
Qy 488 TCATCGTATTCACAGACCTGGGAGTTCTGTATGCGAAGATGCGCCCTTTACGACGTGG 547
Db 6458 TTAATCGTATTCACAGATTTGGGGGTTCTGTGTGCGAAGAAATGGCCCTTTACGATGG 6517
Qy 548 TTTCCACTCTTCTCAGCGCGTGTATGCGCTCTCATACGAGATTCACATACCTCTCTTAAGC 607
Db 6518 TCTCCACCTCTCCTCAGCGCGTGTATGCGCTCTCATACGAGATTCACATACCTCTCTTAAGC 6577
Qy 608 AGCGGCTCGAGTTCTCTGTGTAATACCTTGGAAAGCAAGAAATGCGCTTATGGGCTTCTCAT 667
Db 6578 AGCGGCTCGAGTTCTCTGTGTAATGCTTGGAAAGCAAGAAATGCGCTTATGGGCTTCTCAT 6637
Qy 668 ATGACACCGCTGTTTGTGACTCAACGGTCACTGAGAAATGACATCCGTGTTGAGGAGTCAA 727
Db 6638 ATGACACCGCTGTTTGTGACTCAACGGTCACTGAGAAATGACATCCGTGTTGAGGAGTCAA 6697
Qy 728 TTTACCAATGTTGACTTTGGCCCCCGAAGCTAGACAGGCCATAAGGTGCGTCAAGAGC 787
Db 6698 TCTACCAATGTTGACTTTGGCCCCCGAAGCTAGACAGGCCATAAGGTGCGTCAAGAGC 6757
Qy 788 GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTTGGCGCTATCGCGGT 847
Db 6758 GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTTGGCGCTATCGCGGT 6817
Qy 848 GCCGCGGAGCGGCTGCTGACGACTAGCTGCGGTATACCTCAGCTGCTTGAAG 907
Db 6818 GCCGCGGAGCGGCTGCTGACGACTAGCTGCGGTATACCTCAGCTGCTTGAAG 6877
Qy 908 CCGCTGACGCTGTCGAGCTGCCAAGCTCCAGGACTCCAGCATGCTGCTGTAATGGAGCG 967
Db 6878 CCGCTGACGCTGTCGAGCTGCCAAGCTCCAGGACTCCAGCATGCTGCTGTAATGGAGCG 6937
Qy 968 ACCTTGTGTTATCTGTGAAAGCGCGGAAACCCCAAGAGGAGCGCGCAAGCTTACGAGTCT 1027
Db 6938 ACCTTGTGTTATCTGTGAAAGCGCGGAGACCCCAAGAGGAGCGCGCAAGCTTACGAGTCT 6997
Qy 1028 TCAGGAGGCTAGACTAGGTAAGTCTGCTGCGGCGGAGCGCGGCGGCGGCGGCGGCGGCGG 1087
Db 6998 TCAGGAGGCTAGACTAGGTAAGTCTGCTGCGGCGGAGCGCGGCGGCGGCGGCGGCGGCGG 7057
Qy 1088 ACTTGGAGCTGATAAATCATGTTCTTCTGCAATGTTGCTGCGGAGCGAGCATGCTGCGCA 1147
Db 7058 ACTTGGAGCTGATAAATCATGTTCTTCTGCAATGTTGCTGCGGAGCGAGCATGCTGCGCA 7117
Qy 1148 AAAGGGTGTACTACTCTCACCGGTGACCGCGTCCCGCTTGGCGGGGCTGCGTGGGAGA 1207
Db 7118 AAAGGGTGTACTACTCTCACCGGTGACCGCGTCCCGCTTGGCGGGGCTGCGTGGGAGA 7177
Qy 1208 CAGTAGGCGACTTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGCGGCGGCGGCT 1267
Db 7178 CAGTAGGCGACTTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGCGGCGGCGGCT 7237
Qy 1268 TGTGGGCAAGGATGTTCTGATGACTCTCTTCTTCTCCATCTCTTCTAGGCGGCGGAGCAAC 1327
Db 7238 TGTGGGCAAGGATGTTCTGATGACTCTCTTCTTCTCCATCTCTTCTAGGCGGCGGAGCAAC 7297
Qy 1328 TTGAAAAGCGCTTGAAGCTTGTGAGTCTACGGGCTTGTACTCTCATGAGCGACTTGAACC 1387
Db 7298 TTGAAAAGCGCTTGAAGCTTGTGAGTCTACGGGCTTGTACTCTCATGAGCGACTTGAACC 7357
Qy 1388 TACTCAGATCATTTGAAGCTTGTGAGTCTTGTAGCGGCTTGTACTCTCATGAGCGACTTGAACC 1447
Db 7358 TACTCAGATCATTTGAAGCTTGTGAGTCTTGTAGCGGCTTGTACTCTCATGAGCGACTTGAACC 7417
```

QY 1448 CAGGTGAGATCAATAGGGTGGCTTCATGCCTCAGGAAGCTTGGGTTACCAACCTTGGGAG 1507
DB 7418 CAGGTGAGATCAATAGGGTGGCTTCATGCCTCAGGAAGCTTGGGTTACCAACCTTGGGAG 7477
QY 1508 TCTGAGACATCGGGCCAGAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGAGGGCCG 1567
DB 7478 TCTGAGACATCGGGCCAGAAGTGTCCGCGCTAGCTACTGTCCAGGGGGAGGGCTG 7537
QY 1568 CCAATTTGGCAAGTACCTCTTCAACTGGGCGAGTAAGACCAAGCTTAAACTCACTCAA 1627
DB 7538 CCACTTGGCAAGTACCTCTTCAACTGGGCGAGTAAGACCAAGCTCAAACTCACTCAA 7597
QY 1628 TTCGGCTCGTCCGGCTGGACTTGTCCGCTGGTTCGTTGCTGGCTCAGCGGGGAG 1687
DB 7598 TCCGGCTCGTCCGCTGGAGTTGGAATTAATCCAGCTGGTTCGTTGCTGTTACAGCGGGGAG 7657
QY 1688 ACATATATCAGAGCTGTCTCGTGGCCGACCCCGC 1722
DB 7658 ACATATATCAGAGCTGTCTCGTGGCCGACCCCGC 7692

RESULT 11

US-10-897-648-17
; Sequence 17, Application US/10897648
; Publication No. US20050043266A1
; GENERAL INFORMATION:
; APPLICANT: Jayasena, Sumedha
; APPLICANT: Richardson, Christopher Donald
; TITLE OF INVENTION: SHORT INTERFERING RNA AS AN ANTIVIRAL AGENT FOR HEPATITIS C
; FILE REFERENCE: A-835
; CURRENT APPLICATION NUMBER: US/10/897,648
; CURRENT FILING DATE: 2004-07-22
; PRIOR APPLICATION NUMBER: 60/490,204
; PRIOR FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 7989
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-897-648-17

Query Match 87.8%; Score 1527.8; DB 21; Length 7989;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1598; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 8 GCATGTCAATGTCTTATACATGCAGCGCGCCCTGATCACACCGTGCCTCGGAGGAAA 67
DB 5978 GGTGCTCGATGTCTTACATGGACAGCGGCCCTGATCAGGCCATGCGCTCGGAGGAAA 6037
QY 68 GCAAGTCCCAATCAACGCGCTGAGCAACTCTTCTGCTGCTCAACATTAACCTGTCTATT 127
DB 6038 CCAAGTCCCAATCAATGCACTGAGCAACTTTTCTCGCTCACCACCAACTTGTGCTATG 6097
QY 128 CCACAACTCCCGAGTGAAGCTTGGGCGAGCAAGGTACCTTTGACAGACTGCGAAG 187
DB 6098 CTACAACTCTCGAGCGCAAGCTTGGGCGAGCAAGGTACCTTTGACAGACTGCGAAG 6157
QY 188 TCTTGACGATCATTTACCGGACGTGCTCAAGGAGATGAAGCGAAGCGCTCCACAGTGA 247
DB 6158 TCTTGACGACCATCTACCGGACGTGCTCAAGGAGATGAAGCGAAGCGCTCCACAGTGA 6217
QY 248 AGGCTAAACTGTCTGTGAGGAAGCATGCAAGCTGACGCCCCCGCATTTGGGCCAAAT 307
DB 6218 AGGCTAAACTGTCTGTGAGGAAGCATGTAAGCTGACGCCCCCGCATTTGGGCCAAT 6277
QY 308 CCAAAATTTGGCTATGGGGCAAAAGGAGCTCCGAGGCTTCCAGAGGCGCTTTACCCACA 367
DB 6278 CTAATTTGGCTATGGGGCAAAAGGAGCTCCGGAACCTATCCAGCAAGGCGGTAAACCACA 6337
QY 368 TCCGCTCCGCTGGAGGACTTGTGGAGGACACTGACACCAACCAATTCAGACCCACATCA 427
DB 6338 TCCGCTCCGCTGGAGGACTTGTGGAGGACACTGAGACACCAATTTGACACCCACATCA 6397

QY 428 TGGCAAAAATGAGGTTTTCTCGGTCCAAACACAGAGAAAGGAGGCCGCAACACAGCTCGCC 487
DB 6398 TGGCAAAAATGAGGTTTTCTCGGTCCAAACACAGAGAAAGGAGGCCGCAACACAGCTCGCC 6457
QY 488 TCATCGTATTCCCGAGACCTGGGAGTTGCTGTATGCGAGAAAGATGGCCCTTTACGAGCTGG 547
DB 6458 TTATCGTATTCCCGAGATTTGGGGGTTGCTGTGCGAGAAATGGCCCTTTACGAGTGG 6517
QY 548 TTTCCTCACTCTTCTCAGGCCGTGATGGGCTCTCTCATACGGATTTCCAATATCTCTCTAAGC 607
DB 6518 TCTCCAGCCCTCCCTCAGGCGGTGATGGGCTCTTATACGGATTTCCAATATCTCTCTGGAC 6577
QY 608 AGCGGTGAGTTCCTGGTGAATACCTGGAAAGCAAGAAATGCCCCTATGGGCTTCTCAT 667
DB 6578 AGCGGTGAGTTCCTGGTGAATGCTGGAAAGCAAGAAATGCCCCTATGGGCTTCTCAT 6637
QY 668 ATGACACCCGCTGTTTTGACTCAACGCTCACTGAGAAATGACATCCCTGTTGAGGAGTCAA 727
DB 6638 ATGACACCCGCTGTTTTGACTCAACGCTCACTGAGAAATGACATCCCTGTTGAGGAGTCAA 6697
QY 728 TTTACCAATGTTGCTGACTTGGCCCGCGAAGCTAGACAGGCCATTAAGGTCTCTCACAGAGC 787
DB 6698 TCTACCAATGTTGCTGACTTGGCCCGCGAAGCCAGACAGGCCATTAAGGTCTCTCACAGAGC 6757
QY 788 GGCTCTATGTCCGGGGTCCCATGACTAACTCCAAAGGGCAGAACTCGCGCTATCGCCGGT 847
DB 6758 GGCTTTTACATCGGGGGCCCCCTGACTTAATTTCTAAAGGGCAGAACTCGCGCTATCGCCGGT 6817
QY 848 GCGCGGAGCGCGCTGTGACGACTAGCTGCGGTATATACCTCACTGCTACTTCAAGG 907
DB 6818 GCGCGCGAGCGGTGACTGACGACCACTGCGGTAAATACCTTCACATGTTACTTCAAGG 6877
QY 908 CCGCTCGAGCTGTGAGGCTGCCAAGCTCCAGACTGCGAGCTGCTCGTCAATGAGAGACG 967
DB 6878 CCGCTCGAGCTGTGAGGCTGCCAAGCTCCAGACTGCGAGCTGCTGATGTCGAGAGACG 6937
QY 968 ACCTTGTGCTTATCTGTGAAGCGCGGAAACCCAGAGGACGCGGCAAGCTACAGTCT 1027
DB 6938 ACCTTGTGCTTATCTGTGAAGCGCGGGGACCCAGAGGACGAGGCGAGCTACGGGCT 6997
QY 1028 TCACGAGGCTATGACTAGGTACTCTTGCCCCCTTGGGAGACCCGCCCAACCCGGAATACG 1087
DB 6998 TCACGAGGCTATGACTAGATACTCTGCCCCCCCTGGGAGACCCGCCCAACCCGGAATACG 7057
QY 1088 ACTTGGAGCTGATAACATCGTGTCTCCAAATGTGTCGTCGACAGTGCATCTGGCA 1147
DB 7058 ACTTGGAGTTGATAACATCATGCTCTCCAAATGTGTCGTCGCGCAGATGCATCTGGCA 7117
QY 1148 AAGGGGTGACTACTCTCACCGGTGACCCACCTTCCCTTGGCGGGCTGCGTGGGAGA 1207
DB 7118 AAGGGGTGACTACTCTCACCGGTGACCCACCTTCCCTTGGCGGGCTGCGTGGGAGA 7177
QY 1208 CAGCTAGGCACTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGATGTCGCCCACTT 1267
DB 7178 CAGCTAGACACTCCAGTCAATTTCTTGGCTAGGCAACATCMCATGATATGCCGCCACT 7237
QY 1268 TGTGGGCAAGGATGATTCGATGACTCATTTCTTCTCCATCTCTTAGCCCGAGGAGCAAC 1327
DB 7238 TGTGGGCAAGGATGATCTGATGACTCATTTCTTCTCCATCTCTTAGCTCAGGAACAAC 7297
QY 1328 TTGAAAAGCCCTGGATTTGCGATCTACGGGGCTTGTACTCTCATTTAGCCACTTCAACC 1387
DB 7298 TTGAAAAGCCCTTAGATTTGATGATCTACGGGGCTTGTACTCTCATTTAGGCCACTTGACC 7357
QY 1388 TACCTCAGATCATTTGAACGACTCCATGCTCTTAGCGCATTTTCACTCCATAGTTACTCTC 1447
DB 7358 TACCTCAGATCATTTCAACGACTCCATGCGCTTAGCGCATTTTCACTCCATAGTTACTCTC 7417
QY 1448 CAGGTGAGATCAATAGGGTGGCTTTCATGCTCAGGAAGCTTGGGGTACCAACCTTGGCAG 1507
DB 7418 CAGGTGAGATCAATAGGGTGGCTTTCATGCTCAGGAAGCTTGGGGTACCAACCTTGGCAG 7477

QY 1508 TCTGGAGACATCGGCGCAGAGTGTCCGCTAAGTTACTGTCCAGGGGGGAGGGCG 1567
Db |||||
QY 7478 TCTGGAGACATCGGCGCAGAGTGTCCGCTAAGTTACTGTCCAGGGGGGAGGGCTG 7537
Db |||||
QY 1568 CCATTTGTGGCAAGTACCTCTTCAACTGGGCGAGTAAGGACCAAGCTTAACTCACTCCAA 1627
Db |||||
QY 7538 CCATTTGTGGCAAGTACCTCTTCAACTGGGCGAGTAAGGACCAAGCTTAACTCACTCCAA 7597
Db |||||
QY 1628 TTCCGGCTCGCTCCGCGCTGGACTTGTTCGGCTGGTTCGTTGCTGCTACAGCGGGGAG 1687
Db |||||
QY 7598 TTCCGGCTCGCTCCGCGCTGGACTTGTTCGGCTGGTTCGTTGCTGCTACAGCGGGGAG 7657
Db |||||
QY 1688 ACATATATCACAGCTGTCTCGTCCCGGACCCCGC 1722
Db |||||
QY 7658 ACATATATCACAGCTGTCTCGTCCCGGACCCCGC 7692
Db |||||
RESULT 12
US-10-005-469-1
; Sequence 1, Application US/10005469
; Publication No. US20020155133A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS Pharmaceuticals, Inc.
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY RE
; FILE REFERENCE: 0342/1H395US1
; CURRENT APPLICATION NUMBER: US/10/005,469
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV replicon 1377/NS3-3'UTR
US-10-005-469-1
Query Match 87.8%; Score 1527.8; DB 13; Length 7992;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1598; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 8 GCATGTCAATGTCTTATACATGAGACAGCGCCCTGATACACCGTGCCTGCGGAGGAAA 67
Db |||||
QY 5978 GCTGTCTGATGTCTTACATACATGAGACAGCGCCCTGATACGCGCATGCGCTGCGGAGGAAA 6037
Db |||||
QY 68 GCAAGCTGCCATCAACGGCTGAGCACTCTCTGCTGCTCACCATCACTGCTCTATT 127
Db |||||
QY 6038 CCAAGCTGCCATCAATGCACTGAGCACTCTTCTGCTGCTCACCACCACTTGGTCTATG 6097
Db |||||
QY 128 CCACAACTCCCGCAGTCAAGCTTCCGCGAGAGAGTCACTTTCACAGACTGCAAG 187
Db |||||
QY 6098 CTACAACTCTCGAGCGCAAGCTTCCGCGAGAGAGTCACTTTCACAGACTGCAAG 6157
Db |||||
QY 188 TCTCGACGATCATTAACGGGAGCTGCTCAAGGAGATGAAGCGAGCGCTCCACAGTGA 247
Db |||||
QY 6158 TCTCGACGACCACTACCGGAGCTGCTCAAGGAGATGAAGCGAGCGCTCCACAGTTA 6217
Db |||||
QY 248 AGCTAAACTGTATCTGTAGAGAGCATGCAAGCTGACGCGCCCGGCAATTCGGCCAAAT 307
Db |||||
QY 6218 AGCTAAACTGTATCTCGTGGAGAGAGCTGTAAAGTGAAGCGCCCGCAGATTCGGCCAGAT 6277
Db |||||
QY 308 CCAAAATTTGGCTATGGGCGAAGAGCTCCGAGCTATCCAGAGGCGCGTTAAACACA 367
Db |||||
QY 6278 CTAATAATTTGGCTATGGGCGAAGAGCTCCGAGCCTATCCAGAGGCGCGTTAAACACA 6337
Db |||||
QY 368 TCCGCTCGGTGGAGAGCTTGTGGAGGACACTGACACCAATTCAGACCACTCA 427
Db |||||
QY 6338 TCCGCTCGGTGGAGAGCTTGTGGAGAGACTGCTGGAGAGACTGAGACCAATTCAGACCACTCA 6397
Db |||||
QY 428 TGGCAAAAATGAGGTTTTCTGCTGCCAAACAGAGAGAGGAGCGCGCAAAACCGAGCTCGCC 487
Db |||||

Db |||||
QY 6398 TGGCAAAAATGAGGTTTTCTGCTGCCAAACAGAGAGGAGGGGGCGCAAGCGCTCGCC 6457
QY |||||
QY 488 TCATCGTATTCACAGACCTGGGAGTTCTGTGTATCGAAGATGGCCCTTTACAGACTGG 547
Db |||||
QY 6458 TTATCGTATTCACAGATTTGGGGGTTCTGTGTGGGAGAAATGGCCCTTTACGATGTGG 6517
Db |||||
QY 548 TTTCCACTCTTCTCAGGCGGTGATGGGCTCTCATACGGATTCCAATCTCTCTCTAAGC 607
Db |||||
QY 6518 TCTCACCCCTCTCAGGCGGTGATGGGCTCTCATACGGATTCCAATCTCTCTCTCGAC 6577
Db |||||
QY 608 AGCGGCTCGAGTTCTGTGTGAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 667
Db |||||
QY 6578 AGCGGCTCGAGTTCTGTGTGAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 6637
Db |||||
QY 668 ATGACACCGCTGTGTGTGACTCAACGGTCACTGAGAAATGACATCCGTGTGAGGAGTCAA 727
Db |||||
QY 6638 ATGACACCGCTGTGTGTGACTCAACGGTCACTGAGAAATGACATCCGTGTGAGGAGTCAA 6697
Db |||||
QY 728 TTTACCAATTTGTGACTTTGGCCCGCAGAGCTAGACAGGCCATAGGTCTGCTCACAGAGC 787
Db |||||
QY 6698 TCTACCAATTTGTGACTTTGGCCCGCAGAGCCAGACAGGCCATAGGTCTGCTCACAGAGC 6757
Db |||||
QY 788 GGCTCTATGTGCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGCGGCTTATCGCGGT 847
Db |||||
QY 6758 GGTCTTATCATCGGGGCGCCCTGACTAATTTCTAAAGGCGAGAACTGCGGCTTATCGCGGT 6817
Db |||||
QY 848 GCGGCGAGCGGCTGCTGACGACTAGCTGCGGTGTAATACCTCATCTGCTACTTTGAAG 907
Db |||||
QY 6818 GCGGCGAGCGGTTACTGACGACCACTGCGGTGTAATACCTCATCTGTTGAAAG 6877
Db |||||
QY 908 CCGCTGACAGCTGTGCGAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAATGAGAGC 967
Db |||||
QY 6878 CCGCTGCGGCTGTGAGCTCGAAGCTCCAGGACTGCACGATGCTGCTGTGCGGAGAGC 6937
Db |||||
QY 968 ACCTTGTGTTATCTGTGAAAGCGGGGAAACCCAGAGGAGCGCGCAAGCTTACAGTCT 1027
Db |||||
QY 6938 ACCTTGTGTTATCTGTGAAAGCGGGGAGCCCAAGAGGAGCGGAGCTTACGGGCT 6997
Db |||||
QY 1028 TCACGAGGCTATGACTAGGTACTCTGCGCCCTCTGGGAGCCCGCCCAACCGGAATACG 1087
Db |||||
QY 6998 TCACGAGGCTATGACTAGGTACTCTGCGCCCTCTGGGAGCCCGCCCAACCGGAATACG 7057
Db |||||
QY 1088 ACTTGGAGCTGATTAACATCGTGTCTCAATGTGCTGCGACACGATGCTGCGA 1147
Db |||||
QY 7058 ACTTGGAGTGTATTAACATCATGCTCTCAATGTGCTGAGTGTGCTGCGACGATGCTGCGA 7117
Db |||||
QY 1148 AAAGGGTGTACTACTCTACCCGCTGACCCCGCTGCGGCTGCGGCTGCGGAGG 1207
Db |||||
QY 7118 AAAGGGTGTACTACTCTACCCGCTGACCCCGCTGCGGCTGCGGCTGCGGAGG 7177
Db |||||
QY 1208 CAGCTAGGCGACATCTCCAGTCAACTCTGCTGAGTGAACATCATATGATGCGGCCACTT 1267
Db |||||
QY 7178 CAGCTAGACACATCTCCAGTCAATCTCTGCTGCTGAGTGAACATCATATGATGCGGCCACT 7237
Db |||||
QY 1268 TGTGCGAAGGATGATCTCTGATGACTCACTTCTCTCAATGCTGCTGAGCGGCTGCGGAG 1327
Db |||||
QY 7238 TGTGCGAAGGATGATCTCTGATGACTCACTTCTCTCAATGCTGCTGAGCGGCTGCGGAG 7297
Db |||||
QY 1328 TTGAAAAGCCCTGGATGTCAGATCTACGGGCTTGTACTCTCAATGAGCCACTTGACC 1387
Db |||||
QY 7298 TTGAAAAGCCCTGGATGTCAGATCTACGGGCTTGTACTCTCAATGAGCCACTTGACC 7357
Db |||||
QY 1388 TACTCTAGATCATTTGAACGACTCCATGCTCTTACGCAATTTTCACTCATAGTTACTCTC 1447
Db |||||
QY 7358 TACTCTAGATCATTTGAACGACTCCATGCTCTTACGCAATTTTCACTCATAGTTACTCTC 7417
Db |||||
QY 1448 CAGGTGAGATCAATAGGCTGCTTCTGCTGCTGAGGAGCTTGGGCTACCACTTGGAG 1507
Db |||||
QY 7418 CAGGTGAGATCAATAGGCTGCTTCTGCTGCTGAGGAGCTTGGGCTACCACTTGGAG 7477
Db |||||
QY 1508 TCTGAGACATCGGCGCAGAGTGTGCGGCTAAGTTACTGTGCCAGGGGGGAGGGCG 1567
Db |||||

Db 7478 TCTGGAGACATCGGCGCCAGAAGTGTCCGCGCTAGGCTACTGTCCACGGGGAGGCGCTG 7537

Qy 1568 CCATTGTGGCAAGTACCTCTTCAACTGGGCAGTAAGACCAAGCTTAAACTCACTCAAA 1627

Db 7538 CCATTGTGGCAAGTACCTCTTCAACTGGGCAGTAAGACCAAGCTCAAACTCACTCAAA 7597

Qy 1628 TTCGGCTCGTCCGGCTGGACTTGTCCGCTGGTTGCTGGTTGCTGGCTACAGCGGGGAG 1687

Db 7598 TCCGGCTCGTCCAGTTGGATTATCCAGCTGGTTGGTTGCTGGTTACAGCGGGGAG 7657

Qy 1688 ACATATATCAGAGCTGTCTGTGCGCCGACCCCGC 1722

Db 7658 ACATATATCAGAGCTGTCTGTGCGCCGACCCCGC 7692

RESULT 13

US-10-005-469-2

; Sequence 2, Application US/10005469

; Publication No. US20020155133A1

; GENERAL INFORMATION:

; APPLICANT: ANADYS Pharmaceuticals, Inc.

; APPLICANT: Bichko, Vadim

; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY RE

; FILE REFERENCE: 0342/IH395US1

; CURRENT APPLICATION NUMBER: US/10/005,469

; CURRENT FILING DATE: 2002-04-18

; PRIOR APPLICATION NUMBER: US 60/245,866

; PRIOR FILING DATE: 2000-11-07

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 7992

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: HCV Replicon RNA from cell line HCVR2

US-10-005-469-2

Query Match 87.8%; Score 1527.8; DB 13; Length 7992;

Best Local Similarity 93.2%; Pred. No. 0;

Matches 1598; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Qy 8 GCATGTCAATGTCTATACATGGACAGGCGCCCTGATCACACCGTGGCTCGCGAGGAAA 67

Db 5978 GCTGCTCATGTCTTACACATGGACAGGCGCCCTGATCACGCCATGGCGTGGCGAGGAAA 6037

Qy 68 GCAAGCTGCCATCAACGCGTGTAGCAACTCTTGTCTGCGTCAACATAACCTGGTCTATT 127

Db 6038 CCAAGCTGCCATCAATGCATGTAGCAACTCTTGTCTGCGTCAACATAACCTGGTCTATTG 6097

Qy 128 CCACAACTCCCGAGTGCAGCCCTGCGGCAGAGAAGGTCACTTTTGACAGACTGCAAG 187

Db 6098 CTACAACTCTCGACGCGAGCCCTGCGGCAGAGAAGGTCACTTTTGACAGACTGCGAGG 6157

Qy 188 TCTTGGACGATCATTTACCGGGACGTGTCTCAAGGAGATGAAGGGAAGGCGTCCACAGTGA 247

Db 6158 TCTTGGACGACCACTACCGGGACGTGTCTCAAGGAGATGAAGGGAAGGCGTCCACAGTTA 6217

Qy 248 AGGCTAAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCGCCCGCATTCGGGCCAAT 307

Db 6218 AGGCTAAACTTCTATCTCGTGGAGGAAGCTGTGAAGCTGACGCGCCCGCATTCGGGCCAGAT 6277

Qy 308 CCAAAATTTGGCTATGGGCAAGGACGTCCGAGCGCTATCCAGCAGGCGCTTAACCAACA 367

Db 6278 CTAATTTTGGCTATGGGCAAGGACGTCCGGAACCTATCCAGCAAGGCGCGTTAACCAACA 6337

Qy 368 TCCGCTCCGTGTGGAAGGACTTGTCTGGAGGACACTGTGACACCACTCAATTCAGACCCCATCA 427

Db 6338 TCCGCTCCGTGTGGAAGGACTTGTCTGGAGGACACTGTGACACCACTCAATTCAGACCCCATCA 6397

Qy 428 TGGCAAAAATAGAGTCTTCTGCGTCCAAACAGAGAAAGGAGGCGCGCAACACAGTCTGCC 487

Db 6398 TGGCAAAAATAGAGTCTTCTGCGTCCAAACAGAGAAAGGAGGCGCGCAACACAGTCTGCC 6457

Qy 488 TCATCGTATTCACAGACCTGGAGTTCGTGTATGCGAGAAGATGGCCCTTTACGACGTGG 547

Db 6458 TTATCGTATTCACAGATTTGGGGTTTCGTGTGTGCGAGAAAATGGCCCTTTACGATGTGG 6517

Qy 548 TTTCCACTCTTCTCCTAGGCGGTGATGGGCTCCTCATACGGATTCMAATATCTCTCTTAAGC 607

Db 6518 TCTCCACCCCTCCTCAGGCGGTGATGGGCTCTTCATACGATTTCCAATATCTCTCTCTGGAC 6577

Qy 608 AGCGGTCGAGTTCCTGCTGAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 667

Db 6578 AGCGGTCGAGTTCCTGCTGAATGCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCCAT 6637

Qy 668 ATGACACCCGCTGTTTGTGACTCAACGCTCACTCAGAAATGACATTCCTGTGTGAGGAGTCAA 727

Db 6638 ATGACACCCGCTGTTTGTGACTCAACGCTCACTGAGNAATGACATTCCTGTGTGAGGAGTCAA 6697

Qy 728 TTTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCCCATTAAGGTGCTCTACAGAGC 787

Db 6698 TCTACCAATGTTGTGACTTGGCCCCCGAAGCAGACAGGCCCATTAAGGTGCTCTACAGAGC 6757

Qy 788 GGCTCTATGTCGGGGGTCCCATGACTTAACCTCCAAAGGGCAGAACTGCGGCTATCGCGGT 847

Db 6758 GGCTTTACATCGGGGGCCCCCTGACTTAATTTCTAAAGGGGAGAACTGCGGCTATCGCGGT 6817

Qy 848 GCCGCGGAGCGGCTGTGACGACTAGCTGGGGTAATACCTCTACATGCTACTTCTCAAGG 907

Db 6818 GCCGCGGAGCGGCTGTGACGACTAGCTGGGGTAATACCTCTACATGCTACTTCTCAAGG 6877

Qy 908 CCGCTGCAGCTGTCCAGCTGCCAAGCTCCAGACTGCACGATGCTCGTGAATGGAGAGCG 967

Db 6878 CCGCTGGGCTGTCCAGCTGCCAAGCTCCAGACTGCACGATGCTCGTGAATGGAGAGCG 6937

Qy 968 ACCTTGTGTTATCTGTGAAAGCGCGGAAACCCAGAGGACGCGGCAAGCCTACAGTCT 1027

Db 6938 ACCTTGTGTTATCTGTGAAAGCGCGGGAACCCAGAGGACGAGGCGAGCCTACGCGCT 6997

Qy 1028 TCAGGAGGCTATGACTAGTACTCTGCCCGCTGGGACCGCGCCCAACCGGATACG 1087

Db 6998 TCAGGAGGCTATGACTAGTACTCTGCCCGCTGGGACCGCGCCCAACCGGATACG 7057

Qy 1088 ACTTGGAGCTGATAACATCGTGTCTCCAAATGTGTGCGGTGCGACACGATGCAATCTGGCA 1147

Db 7058 ACTTGGAGTTGATAAATCATGCTCTCTCAATGTGTGAGTGGCGACAGATGCAATCTGGCA 7117

Qy 1148 AAAGGTTGTAACCTCACCCGTGACCCCAACCGCTCCCTTGGCGGGCTGCGTGGGAGA 1207

Db 7118 AAAGGTTGTAACCTCACCCGTGACCCCAACCGCTCCCTTGGCGGGCTGCGTGGGAGA 7177

Qy 1208 CAGCTAGGCACACTCCAGTCAACTCTCTGGGTAGGCAACATCATCTGATGTATGCGGCCACTT 1267

Db 7178 CAGCTAGGCACACTCCAGTCAACTCTCTGGGTAGGCAACATCATCTGATGTATGCGGCCACTT 7237

Qy 1268 TGTGGCGCAAGATGATTTCTGATGACTCACTTCTTCTCCATCTCTTAGCCAGAGCAAC 1327

Db 7238 TGTGGCGCAAGATGATTTCTGATGACTCACTTCTTCTCCATCTCTTAGCCAGAGCAAC 7297

Qy 1328 TTGAAAAGCCCTGGGATTTGCAGATCTACGGGCTTGTACTCTCAATTTAGGCCACTTTGACC 1387

Db 7298 TTGAAAAGCCCTAGATTTGCAGATCTACGGGCTTGTACTCTCAATTTAGGCCACTTTGACC 7357

Qy 1388 TACCTCAGATCAATTGAACCGACTCCATGGCTTTTAGCGCATTTTCACTCCATPACTCTC 1447

Db 7358 TACCTCAGATCAATTGAACCGACTCCATGGCTTTTAGCGCATTTTCACTCCATPACTCTC 7417

Qy 1448 CAGGTGAGATCAATAGGCTGGCTTCATGCTCTAGGAGAGCTTGGGGTACCCCTCTGGGAG 1507

Db 7418 CAGGTGAGATCAATAGGCTGGCTTCATGCTCTAGGAGAGCTTGGGGTACCCCTCTGGGAG 7477

Qy 1508 TCTGGAGACATCGGCGCCAGAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGCGCG 1567

Db 7478 TCTGGAGACATCGGCGCCAGAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGCGCTG 7537

Db 7538 CCACTTGTGGCAAGTACCTCTTCAACTGGGCAGTAAGGACCAAGCTCAAACCTCACTCCAA 7597
Qy 1628 TTCCGGCTCGCTCCGGCTGGACTTGTCCGGCTGGTTGCTGGCTACAGCGGGGAG 1687
Db 7598 TCCCGCTCGCTCCAGCTGGATTATCCAGCTGGTTGCTGCTTACAGCGGGGAG 7657
Qy 1688 ACATATATCAGAGCTGTCTCGTCCCGACCCCGC 1722
Db 7658 ACATATATCAGAGCTGTCTCGTCCCGACCCCGC 7692

RESULT 15

US-10-005-469-5
; Sequence 5, Application US/10005469
; Publication No. US20020155133A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS Pharmaceuticals, Inc.
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPLICATION
; FILE REFERENCE: 0342/IH395U51
; CURRENT APPLICATION NUMBER: US/10/005,469
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Replicon from cell line HCVp22

Query Match 87.8%; Score 1527.8; DB 13; Length 7992;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1598; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Qy 8 GCATGTCAATGTCTATACATGGACAGCGCCCTGATCAGACCGCTGGCTCGCGAGGAAA 67
Db 5978 GCTGCTCATGTCTTACACATGGACAGCGCCCTGATCAGCCCATGGCTGCGGAGGAAA 6037
Qy 68 GCAAGCTGCCATCAACGCGCTGAGCAACTCTTGTCTGCGTCACCATAACTGTGTATT 127
Db 6038 CCAAGCTGCCATCAATGCATCTGAGCAACTCTTGTCTGCGTCACCATAACTTGTGTATT 6097
Qy 128 CCACAACTCCCGAGTGCAGAGCTGCGGAGAGAGAGGTCAACCTTTGACAGACTGCAAG 187
Db 6098 CTACAACTCTCGAGCGAGAGCTGCGGAGAGAGAGGTCAACCTTTGACAGACTGCAAG 6157
Qy 188 TCCTGGACGATCATTTACCGGACGCTGCTCAAGGAGATGAAGGAGAGCGTCCACAGTGA 247
Db 6158 TCCTGGACGACCATACCGGACGCTGCTCAAGGAGATGAAGGAGAGCGTCCACAGTGA 6217
Qy 248 AGGCTAAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCCCGCCCGCATTTGCGGCAAT 307
Db 6218 AGGCTAAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCCCGCCCGCATTTGCGGCAAT 6277
Qy 308 CCNAATTTGGCTATGGGCAAGGAGCTGCGGAGCTATCCAGAGAGCGGCTTAACCACTCA 367
Db 6278 CTAATTTGGCTATGGGCAAGGAGCTGCGGAGCTATCCAGAGAGCGGCTTAACCACTCA 6337
Qy 368 TCCGCTCCGCTGGAAGGACTTGTCTGGAGGACACTGACACACCAATTCAGACACCACTCA 427
Db 6338 TCCGCTCCGCTGGAAGGACTTGTCTGGAGGACACTGAGACACCAATTCAGACACCACTCA 6397
Qy 428 TGGCAAAAATAGGTTTCTGCTGCCAAACAGAGAGAGGCGCGCAACACAGCTCGCC 487
Db 6398 TGGCAAAAATAGGTTTCTGCTGCCAAACAGAGAGAGGCGCGCAACACAGCTCGCC 6457
Qy 488 TCATCGTATCCAGACCTGGAGTTCTGTATGCGAGAGAGTGGCCCTTTACGACGTGG 547
Db 6458 TTATCGTATCCAGAGTTTGGGGGTTCTGTGTGCGAGAGAAATGGCCCTTTACGATGTGG 6517

Qy 548 TTTCACACTCTTCTCCTCAGGCGGTGATGGGCTCTCTATACCGAATTCGAATCTCTCTTAAGC 607
Db 6518 TCTCCACCCCTCCTCAGGCGGTGATGGGCTCTCTATACCGAATTCGAATCTCTCTCTGGAC 6577
Qy 608 AGCGGCTCAGTTCCTGGTGAATACCTGGAAACGAAGAAATGCCCTATGGGCTTCTCAT 667
Db 6578 AGCGGCTCAGTTCCTGGTGAATGCTGGAAACGAAGAAATGCCCTATGGGCTTCTCAT 6637
Qy 668 ATGACACCCGCTGTTTGTGACTCAACCGGTCACTGAGAAATGCAATCCGTGTTGAGAGTCAA 727
Db 6638 ATGACACCCGCTGTTTGTGACTCAACCGGTCACTGAGAAATGCAATCCGTGTTGAGAGTCAA 6697
Qy 728 TTATACCAATTTGTGACTTTGGCCCCCGAAGCTAGACAGGCCCATTAAGGTCTGCTCACAGAGC 787
Db 6698 TCTACCAATTTGTGACTTTGGCCCCCGAAGCTAGACAGGCCCATTAAGGTCTGCTCACAGAGC 6757
Qy 788 GGCTCTATGTGGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGCGGCTATTCGCGGCT 847
Db 6758 GGCTTTACATCGGGGCGCCCTGACTTAATTTCTAAAGGCGAGAACTGCGGCTATTCGCGGCT 6817
Qy 848 GCGCGCGAGCGCGCTGCTGACGACTAGCTGGGTAAATACCTCTACATCTGACTTCTCAAGS 907
Db 6818 GCGCGCGAGCGCGTGTACTGACGACCAAGCTGCGGTAAATACCTCTACATCTGACTTCTCAAGS 6877
Qy 908 CCGCTGCAGCTGTGAGCTGCCAAGCTCCAGACTGCGACGATGCTCGTGAATGGAGAGC 967
Db 6878 CCGCTGGGCTGTGAGCTGCGAAGCTCCAGACTGCGAGCTGCGAGTCTGCTATGGGAGAGC 6937
Qy 968 ACCTTGTCTTATCTGTGAAGCGCGGAAACCCAGAGGACGCGGCAAGCCTACAGTCT 1027
Db 6938 ACCTTGTCTTATCTGTGAAGCGCGGGAACCCAGAGGACGAGGCGAGCCTACGCGCT 6997
Qy 1028 TCACGAGGCTATGACTAGTACTTGGCCCCCTCTGGGAGACCGGCCCAACCGGAATACG 1087
Db 6998 TCACGAGGCTATGACTAGTACTTGGCCCCCTCTGGGAGACCGGCCCAACCGGAATACG 7057
Qy 1088 ACTTGGAGCTGATAACATCGTGTTCCTCAATGTGTGCGGTGCGACACAGATGATCTGGCA 1147
Db 7058 ACTTGGAGTTGATAACATCATGCTCTCTCAATGTGTGCGGTGCGACAGATGATCTGGCA 7117
Qy 1148 AAAGGCTGTACTACCTCACCGGTGACCCCAAGCTGCGGCTGCGGCGGCTGCGTGGAGAG 1207
Db 7118 AAAGGCTGTACTATCTCACCGGTGACCCCAAGCTGCGGCTGCGGCGGCTGCGTGGAGAG 7177
Qy 1208 CAGCTAGGACACTCCAGTCAACTCTCTGGGTAGGCAACATCATCATGTATGCGGCCCACTT 1267
Db 7178 CAGCTAGACACTCCAGTCAATTCCTGGCTAGGCAACATCATCATGTATGCGGCCCACTT 7237
Qy 1268 TGTGGGCAAGGATGATTCGATGACTCACTTCTTCTCCATCTCTTAGCCCGAGAGCAAC 1327
Db 7238 TGTGGGCAAGGATGATTCCTGATGACTCATTTCTCTCCATCTCTTAGCTCAGGAACTAC 7297
Qy 1328 TTCAAAAGCCCTGGATGTGCAGATCTAGGGGCTTGTACTCCATTTAGGCCACTTTGACC 1387
Db 7298 TTGAAAAGCCCTTAGATGTGCAGATCTAGGGGCTTGTACTCCATTTAGGCCACTTTGACC 7357
Qy 1388 TACTCAGATCATTTGAAACGACTCCATGGTCTTAGCGCATTTTCACTCCATAGTACTCTC 1447
Db 7358 TACTCAGATCATTTCAACGACTCCATGGCTTAGCGCATTTTCACTCCATAGTACTCTC 7417
Qy 1448 CAGGTGAGATCAATAGGGTGGCTTCATGCGCTCAGGAAGCTTGGGGTACCACTTGGGAG 1507
Db 7418 CAGGTGAGATCAATAGGGTGGCTTCATGCGCTCAGGAACCTTGGGGTACCGCCCTTGGCAG 7477
Qy 1508 TCTGGGACATCGGCGCAGAGGTGCGGCTCAAGTTACTGTCTCCAGGGGGGAGGGCGC 1567
Db 7478 TCTGGGACATCGGCGCAGAGGTGCGGCTCAAGTTACTGTCTCCAGGGGGGAGGGCTG 7537
Qy 1568 CCATTTGTGGCAAGTACTCTTCAACTGGGCACTAAGGACCAAGCTTAAACTCACTCCAA 1627
Db 7538 CCATTTGTGGCAAGTACTCTTCAACTGGGCACTAAGGACCAAGCTTAAACTCACTCCAA 7597

```
Qy 1628 TTCGGCTGGTCCCGCTGGACTTGTCCGGCTGGTTCGTTGCTGGCTACAGCGGGGAG 1687
Db 7598 TCCCGCTGGTCCCGCTGGACTTGTCCGGCTGGTTCGTTGCTGGCTACAGCGGGGAG 7657
Qy 1688 ACATATATCACAGCCTGTCTCGTGCCCGACCCCGC 1722
Db 7658 ACATATATCACAGCCTGTCTCGTGCCCGACCCCGC 7692
```

Search completed: October 3, 2005, 23:09:02
Job time : 1231.79 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 13:25:15 ; Search time 298.103 Seconds
(without alignments)
9550.788 Million cell updates/sec

Title: US-10-712-479-3

Perfect score: 1740
Sequence: 1 atggctagcatgtcaatgtc.....gccatcaccatcacatcac 1740

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1551.8	89.2	9472	4	US-08-150-204E-96
2	1547	88.9	9595	3	US-09-014-416-4
3	1539.6	88.5	2991	1	US-08-324-977-49
4	1539.6	88.5	2991	2	US-08-384-616-49
5	1539.6	88.5	2991	2	US-08-904-686A-49
6	1539.6	88.5	2991	3	US-09-315-850-49
7	1539.6	88.5	7863	1	US-08-324-977-35
8	1539.6	88.5	7863	2	US-08-384-616-35
9	1539.6	88.5	7863	2	US-08-904-686A-35
10	1539.6	88.5	7863	3	US-09-315-850-35
11	1539.6	88.5	7917	1	US-08-324-977-31
12	1539.6	88.5	7917	2	US-08-384-616-31
13	1539.6	88.5	7917	2	US-08-904-686A-31
14	1539.6	88.5	7917	3	US-09-315-850-31
15	1539.6	88.5	9030	1	US-08-324-977-13
16	1539.6	88.5	9030	2	US-08-384-616-13
17	1539.6	88.5	9030	2	US-08-904-686A-13
18	1539.6	88.5	9030	3	US-09-315-850-13
19	1539.6	88.5	9416	1	US-08-324-977-1
20	1539.6	88.5	9416	2	US-08-384-616-1
21	1539.6	88.5	9416	2	US-08-904-686A-1
22	1539.6	88.5	9416	3	US-09-315-850-1
23	1539.6	88.5	9416	3	US-08-823-895A-27
24	1527.8	87.8	7989	4	US-09-539-601-10
25	1527.8	87.8	8001	4	US-09-539-601-7
26	1527.8	87.8	8001	4	US-09-539-601-22
27	1527.8	87.8	8637	4	US-09-539-601-4

ALIGNMENTS

RESULT 1

US-08-150-204E-96
; Sequence 96, Application US/08150204E
; Patent No. 6538126
; GENERAL INFORMATION:

APPLICANT: CHO, Joong Myung
LEE, Yong Beom
PARK, Young Woo
LIM, Kook Jin
CHOI, Deog Young
SO, Hong Seob
KIM, Chun Hyung
KIM, Sung Taek
YANG, Jae Young

TITLE OF INVENTION: HEPATITIS C DIAGNOSTICS AND VACCINES
NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESS:

ADDRESSER: YANG, Jae Young
STREET: 386-1, Doryong-dong, Yuseong-gu
CITY: Daejeon
STATE: Daejeon
COUNTRY: Republic of Korea
ZIP: 305-340

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5inch, 1.44MB storage
COMPUTER: IBM PC/pentium
OPERATING SYSTEM: Windows
SOFTWARE: Microsoft Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/150,204E
FILING DATE: 20-Apr-1994
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: KR 91-9510
FILING DATE: 10-JUN-1991
APPLICATION NUMBER: KR 91-13601
FILING DATE: 6-AUG-1991

ATTORNEY/AGENT INFORMATION:
NAME: Shahan Islam, Esq.
REGISTRATION NUMBER: 32,507
REFERENCE/DOCKET NUMBER: 2695/FLX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 940-8564
TELEFAX: (212) 940-8776

INFORMATION FOR SEQ ID NO: 96
SEQUENCE CHARACTERISTICS:
LENGTH: 9472 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

Sequence 13, Appl
Sequence 1, Appl
Sequence 25, Appl
Sequence 28, Appl
Sequence 6, Appl
Sequence 7, Appl
Sequence 24, Appl
Sequence 25, Appl
Sequence 1, Appl
Sequence 2, Appl
Sequence 31, Appl
Sequence 16, Appl
Sequence 4, Appl
Sequence 5, Appl
Sequence 19, Appl
Sequence 6, Appl
Sequence 1, Appl
Sequence 11, Appl

28 1527.8 87.8 8649 4 US-09-539-601-13
29 1527.8 87.8 11076 4 US-09-539-601-1
30 1527.8 87.8 11076 4 US-09-539-601-25
31 1526.2 87.7 8001 4 US-09-539-601-28
32 1526.2 87.7 8638 4 US-10-029-907-6
33 1526.2 87.7 8638 4 US-10-029-907-7
34 1526.2 87.7 8638 4 US-10-029-907-24
35 1526.2 87.7 8638 4 US-10-029-907-25
36 1526.2 87.7 8639 4 US-10-029-907-1
37 1526.2 87.7 8642 4 US-10-029-907-2
38 1526.2 87.7 11076 4 US-09-539-601-31
39 1524.6 87.6 8001 4 US-09-539-601-16
40 1524.6 87.6 8643 4 US-10-029-907-4
41 1524.6 87.6 8648 4 US-10-029-907-5
42 1524.6 87.6 11076 4 US-09-539-601-19
43 1523.6 87.6 9413 4 US-09-827-688-6
44 1522.8 87.5 1773 4 US-09-720-095A-1
45 1519.8 87.3 1782 3 US-09-597-877-11

MOLECULE TYPE: DNA									
FEATURE:									
OTHER INFORMATION: RHCY-LBCL, Fig. 2									
SEQUENCE DESCRIPTION: SEQ ID NO: 96									
US-08-150-204E-96									
Query Match 89.2%; Score 1551.8; DB 4; Length 9472;									
Best Local Similarity 94.1%; Pred. No. 0;									
Matches 1613; Conservative 0; Mismatches 102; Indels 0; Gaps 0;									
QY	8	GCATGTCATGTCCTATATACATGCGAGCGCCCTGATCACACCGTGGCTCGGAGGAAA	67						
DB	7595	GCTGTTCATGTCCTACACATGGACAGGCGGCCCTGATCACGCCATGGCTCGGAGGAAA	7654						
QY	68	GCAAGCTGCCATCAACCGCTGAGCAACTCTCTTGCTGCTCACCAATACCTGGTCTATT	127						
DB	7655	GCAAGTTGCCATCAACCGTTGAGCAATCTCTTGCTAGGTCACCAACATGGTCTATG	7714						
QY	128	CCACAACATCCCGCAGTGCAGCCCTGGCGAGAGAGAGGTCACTTTGACAGACTTGCAG	187						
DB	7715	CTACAACATCCCGCAGCGAGCCCTGGCGAGAGAGGTCACTTTGACAGACTTGCAG	7774						
QY	188	TCCTGGACGATCATTTACCGGAGCTGCTCAAGGAGATGAAGCGGAGGCTCCACAGTGA	247						
DB	7775	TCCTGGACGACCACTACCGGAGCTGCTTAAGGAGATGAAGCGGAGGCTCCACAGTTA	7834						
QY	248	AGGCTAAATGCTATCTGTAGAAGAGCATGCAAGCTGACGCGCCCGCATTCGGCCAAAT	307						
DB	7835	AGGCTAAATCTTATCTGTAGAAGAGCTGCTTAAGGAGATGAAGCGGAGGCTCCACAGTTA	7894						
QY	308	CCAAATTTGGCTATGGGCAAGAGGAGCTCCGAGGCTATCCAGAGGCGCGTTAAACCA	367						
DB	7895	CCAAATTTGGCTATGGGCGAGAGGAGCTCCGAGGCTATCCAGAGGCGCGTTAAACCA	7954						
QY	368	TCGGTCCGCTGGAGAGACTTTCCTGGAGGACATTCGACACACCAATTCAGACCACTCA	427						
DB	7955	TCGGTCCGCTGGAGAGACTTTCCTGGAGGAGCTTCGACACCAATTCAGACCACTCA	8014						
QY	428	TGSCAAAAATAGAGTTTCTGGTCCACACAGAGAGAGGCGGCAACCCAGCTCCG	487						
DB	8015	TGSCAAAAATAGAGTTTCTGTGTCCAAACAGAGAGAGGCGGCGCAACCCAGCTCCG	8074						
QY	488	TCACTCGTATTCACAGACCTGGAGTTCTGTGTATGCGAGAGATGGCCCTTTACGACGTGG	547						
DB	8075	TTATCGTGTCCAGATCTGGAGTTCTGTGTATGCGAGAGATGGCCCTTTATGAGTGG	8134						
QY	548	TTTCCACTCTTCTCAGGCGGTGATGGGCTCTCTATACCGATTCCTCAATCTCTCTAAGC	607						
DB	8135	TTCTCCACCTTCTCAGGCGGTGATGGGCTCTCTATACCGATTCCTCAATCTCTCTAAGC	8194						
QY	608	AGCGGTTCGAGTTCTCGTGAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT	667						
DB	8195	AGCGGTTCGAGTTCTCGTGAATACCTGGAAATCAAGAAATGCCCTATGGGCTTCTCAT	8254						
QY	668	ATGACACCCGCTGTTTGACTCAACCGTCACTGAGATGACATCCGTTGTAGGAGTCAA	727						
DB	8255	ATGACACCCGCTGTTTGACTCAACCGTCACTGAGATGACATCCGTTGTAGGAGTCAA	8314						
QY	728	TTTACCAATGTTGACTTGGCCCCCGAGCTAGACAGGCTAAGCTCGCTCACAGAGC	787						
DB	8315	TTTACCAATGTTGACTTGGCCCCCGAGCTAAGCTGAGCTAAGCTCGCTCACAGAGC	8374						
QY	788	GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTCGGCTATCGCGGT	847						
DB	8375	GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTCGGCTATCGCGGT	8434						
QY	848	GCGCGGAGGCGGCTGCTGACGATAGCTGCGGTAATACCTCCATCATGCTACTTGAAG	907						
DB	8435	GCGCGGAGGCGGCTGCTGACGATAGCTGCGGTAATACCTCCATCATGCTACTTGAAG	8494						
QY	908	CCGCTGCAGCTGTCGAGCTGCCAAGCTCCAGGACTCCAGATGCTCGTCAATGGAGACG	967						
DB	8495	CCACTGCGGCTGTCGAGCTGCCAAGCTCCGAGCTCCGAGTCTCGTGAACGGAGCG	8554						

QY	968	ACCTTCTGTTTCTCTGAAAGCGCGGAAACCAAGAGGACGCGGCAAGCTACGAGTCT	1027						
DB	8555	ACCTTCTGTTTCTCTGAAAGCGCGGAAACCAAGAGGATGGCGGAGCTACGAGTCT	8614						
QY	1028	TCAGGAGGCTATGACTAGGTACTCTGCCCCCTCTGGGACCGCCGCCAACCGGAATACG	1087						
DB	8615	TCAGGAGGCTATGACTAGGTACTCTGCCCCCTCTGGGACCGCCGCCAACCGGAATACG	8674						
QY	1088	ACTTGAGCTGATAAACATCGTGTTCCTCCAAATGTGCGGTCCGACACGATGATCTGGCA	1147						
DB	8675	ACTTGAGTGTGATAACATCATGTTCCTCCAATGTGCGGTCCGACACGATGATCTGGTA	8734						
QY	1148	AAAGGTGTACTACTCACCGGTACCCACCGCTCCCTTGGCGGGCTGCGTGGGAGA	1207						
DB	8735	AAAGGTGTACTACTCACCGGTACCCACCGCTCCCTTGGCGGGCTGCGTGGGAGA	8794						
QY	1208	CAGTAGGCACACTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGCCCACTT	1367						
DB	8795	CAGTAGACACACTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGCCCACTT	8854						
QY	1268	TGTGGCAAGGATGATTCATGATGACTCTTCTTCTCCATCTTCTAGCCAGGAGCAAC	1327						
DB	8855	TATGGCAAGGATGATTCATGATGACTCTTCTTCTCCATCTTCTAGCTCAGGAGCAAC	8914						
QY	1328	TTGAAAAGCCCTGGATTCAGATCTACGGGCTTGTACTCCATTTAGCCACTTTGACC	1387						
DB	8915	TTGAAAAGCCCTTAGATTCAGATCTACGGGCTTGTACTCCATTTGAACCACTTTGATC	8974						
QY	1388	TACCTCAGATCANTGAACGACTCCATGCTCTTAGCGCATTTCACTCCATAGTTACTCTC	1447						
DB	8975	TACCTCAGATCANTGAACGACTCCATGCTCTTAGCGCATTTCACTCCATAGTTACTCTC	9034						
QY	1448	CAGGTGAGATCAATAGGGTGGCTTCATGCTCTAGGAAAGTTGGGGTACCACTTGGCG	1507						
DB	9035	CAGGTGAGATCAATAGGGTGGCTTCATGCTCTAGGAAAGTTGGGGTACCACTTGGCG	9094						
QY	1508	TCTGGAGACATCGGGCCAGAGTCTCCGCGCTAAGTTACTGTCCAGGGGGAGGGCCG	1567						
DB	9095	CCTGGAGACATCGGGCCAGAGTCTCCGCGCTAAGTTACTGTCCAGGGGGAGGGCCG	9154						
QY	1568	CCATTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCAA	1627						
DB	9155	CCATTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCAA	9214						
QY	1628	TTCCGGCTGCTCCCGGCTGGACTTGTCCGGCTGCTTGGTGGCTTACAGCGGGGAG	1687						
DB	9215	TCCAGACCGGCTCCGGTTGAGCTTGTCCGGCTGCTTGGTGGCTTACAGCGGGGAG	9277						
QY	1688	ACATATATCACAGCTGCTCTGTCGCCGAGACCCCGC	1722						
DB	9275	ACATATATCACAGCTGCTCTGTCGCCGAGACCCCGC	9309						

RESULT 2

US-09-014-416-4
; Sequence 4, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; EARLIER FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4

```
; LENGTH: 9595
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-014-416-4

Query Match      88.9%; Score 1547; DB 3; Length 9595;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1610; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 8 GCATGTCAATGTCTTATACATGACAGAGCGCCCTGATCACACCGTGGCGTGGAGGAAA 67
DB 7594 GCTGCTCAATGTCTTATACATGTCGACAGAGCGCCCTGATCAGCCATCGCGTGGCGAGGAAA 7653
QY 68 CAAAGTGGCCATCAACGCGCTGAGCAACTCTTGTGCGTCAACATCAACCTGTCTATT 127
DB 7654 GTAAGTGGCCATCAACCGCTTGTGAGCAACTCTTGTGCGTCAACATCAACCTGTCTATT 127
QY 128 CCACAACATCCCGCAGTCAACCGCTCGGACAGAGAGGTACCTTTGACAGACTGCAAG 187
DB 7714 CCACAACATCCCGCAGTCAACCGCTCGGACAGAGAGGTACCTTTGACAGATTGCAAG 7773
QY 188 TCTCGGACGATCATTCACGGGACGTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA 247
DB 7774 TCTCGGATGATCATTCACGGGACGTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA 7833
QY 248 AGGCTAAATGCTATCTGTAGAGAGCATGCAAGCTGACGCGCCCGCAITTCGGCCAAAT 307
DB 7834 AGGCTAAATGCTATCTATAGAGAGGCTGCAAGCTGACGCGCCCGCAITTCGGCCAAAT 7893
QY 308 CCAATTTGGCTATGGGCGAAGGACGTCCGAGGCTATCCAGAGGCGCGTTAAACACA 367
DB 7894 CCAATTTGGCTATGGGCGAAGGACGTCCGGAACCTATCCAGAGGCGCGTTAAACACA 7953
QY 368 TCCGCTCGGTGGAGGACACTGCTGGAGGACACTGACACACCAATTGACACCACTCA 427
DB 7954 TCCGCTCGGTGGAGGACACTGCTGGAGGACACTGACACACCAATTGACACCACTCA 8013
QY 428 TGGCAAAATAGAGTTTCTCGTCCAAACAGAGAGGAGCGCGGAAACCACTCGCC 487
DB 8014 TGGCAAAATAGAGTTTCTCGTCCAAACAGAGAGGAGCGCGGAAACCACTCGCC 8073
QY 488 TCATCGTATCCAGACTGGAGTTCGTGTATCGAGAGATGGCGCTTTTACAGCTGG 547
DB 8074 TTATCGTATCCAGACTGGAGTTCGTGTATCGAGAGATGGCGCTTTTACAGCTGG 8133
QY 548 TTTCACATCTTCCTCAGCGCGTATGGGCTCCTCATACGATTCCTTCTCTTAAGC 607
DB 8134 TCTCACCTTCTCCTCAGCGCGTATGGGCTCCTCATACGATTCCTTCTCTTAAGC 8193
QY 608 AGCGGTCGAGTTCCTCGTGTATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCAT 667
DB 8194 AGCGGTCGAGTTCCTCGTGTATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCAT 8253
QY 668 ATGACACCGCTGTTTGTGACTCAACCGTCTCATGAGATGATCCCGTGTGAGGCTCAA 727
DB 8254 ATGACACCGCTGTTTGTGACTCAACCGTCTCATGAGATGATCCCGTGTGAGGCTCAA 8313
QY 728 TTACCAATGTTGTGACTTGGCCCCCGAAGCTGACAGGCGCATAAGGTCGCTCACAGAGC 787
DB 8314 TTACCAATGTTGTGACTTGGCCCCCGAAGCTGACAGGCGCATAAGGTCGCTCACAGAGC 8373
QY 788 GGCTCTATGTGGGGTCCCATGATTAACCTCAAGAGGCGAGAACTGCGGCTATCGCGGT 847
DB 8374 GGCTCTATGTGGGGTCCCATGATTAACCTCAAGAGGCGAGAACTGCGGCTATCGCGGT 8433
QY 848 GCCGCGAGCGGCTGCTGACGACTAGCTGCGGTAATACCTTCATGCTTCTTGAAG 907
DB 8434 GCCGCGAGCGGCTGCTGACGACTAGCTGCGGTAATACCTTCATGCTTCTTGAAG 8493
QY 908 CCCTGTCAGCTGCTGAGCTGCGCAAGCTCCAGAGCTGCAAGTCTGCTGAATGGAGAGC 967
DB 8494 CCCTGTCAGCTGCTGAGCTGCGCAAGCTCCAGAGCTGCAAGTCTGCTGAATGGAGAGC 9553
```

```
QY 968 ACCTTGTGTTATCTGTGAAAGCGCGGAAACCCAGAGAGAGCGCAAGCCTACGAGTCT 1027
DB 8554 ACCTTGTGTTATCTGTGAAAGCGCGGAAACCCAGAGAGATGCGCGGCCCTACGAGCCT 8613
QY 1028 TCACGAGAGCTATGACTAGGTACTCTGCCCCCTTGGGAGACCGCCCCCAACCGGAATACG 1087
DB 8614 TCACGAGAGCTATGACTAGGTATTCGCGCCCCCGGGGATCCGCCCCCAACCGGAATACG 8673
QY 1088 ACTTGGAGCTGATTAACATCGTTCTTCCATGATGTGCGTGCACACAGATGATCTGCA 1147
DB 8674 ACTTGGAGCTGATTAACATCATGTTCTTCCATGATGTGCGTGCACACAGATGATCTGCA 8733
QY 1148 AAAAGGTGTACTCTCACCGTGACCCCAACCGTCTGGCTAGGCAACATCATGATGCGGCCACTT 1267
DB 8734 AAAAGGTGTACTCTCACCGTGACCCCAACCGTCTGGCTAGGCAACATCATGATGCGGCCACTT 8793
QY 1268 CAGCTAGGACACTCCAGTCAACTCTGGCTAGGCAACATCATGATGCGGCCACTT 1267
DB 8794 CAGCTAGGACACTCCAGTCAACTCTGGCTAGGCAACATCATGATGCGGCCACTT 8853
QY 1268 TGTGGGCAAGATGATTTCTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1327
DB 8854 TATGGGCAAGGATGATTTCTGATGACTCACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8913
QY 1328 TTGAAAAGCCCTCGGATTTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1387
DB 8914 TTGAAAAGCCCTCGGATTTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8973
QY 1388 TACCTCAGATCATTTGAACGACTCCATGCTCTTAGGCAATTTTCTTCTTCTTCTTCTTCTTCTTCT 1447
DB 8974 TACCTCAGATCATTTGAACGACTCCATGCTCTTAGGCAATTTTCTTCTTCTTCTTCTTCTTCTTCT 9033
QY 1448 CAGGTGAGATCAATAGGCTGGCTTCTGCTCTCAGGAAAGCTTGGGCTACCACTTCTTCTTCTTCT 1507
DB 9034 CAGGTGAGATCAATAGGCTGGCTTCTGCTCTCAGGAAAGCTTGGGCTACCACTTCTTCTTCTTCTTCT 9093
QY 1508 TCTGGAGACATCGGCGCAGAAAGTGTGCGGCTAAGTTACTTGTCTTCCAGGGGGGAGGCGCG 1567
DB 9094 CCTGGAGACATCGGCGCAGAAAGTGTGCGGCTAAGTTACTTGTCTTCCAGGGGGGAGGCGCG 9153
QY 1568 CCATTTGTGCAAGTACCTTTCATCTGAGGCAAGCTTCAAGCTGAGGCAAGCTTAAACTCACTCAA 1627
DB 9154 CCATTTGTGCAAGTACCTTTCATCTGAGGCAAGCTTCAAGCTGAGGCAAGCTTAAACTCACTCAA 9213
QY 1628 TTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1722
DB 9214 TTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1687
QY 1688 ACATATATCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9273
DB 9274 ACATATATCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9308
```

```
RESULT 3
US-08-324-977-49
; Sequence 49, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
```

ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2991 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..2991
OTHER INFORMATION: /note: "sequence = 6372 - 9362 of
OTHER INFORMATION: SEQ ID NO: 1"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2991

US-08-324-977-49

Query Match 88.5%; Score 1539.6; DB 1; Length 2991;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1605; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 8 GCATGTCAATGTCTATACATGACACAGCGCCCTGATCACCAGCGTGGCTGGGAGGAA 67
Db 1214 GCTGCTCAATGTCTATACATGACACAGCGCCCTTGATCAGCCATGGCTGGGAGGAA 1273

Qy 68 GCAAGCTGCCCATCAACGCGCTGAGCAACTCTTGCTGGCTACCAATAACTGGTCTATT 127
Db 1274 GCAAGCTGCCCATCAACGCGCTGAGCAACTCTTGCTGGCCCAACCAATACATGTTTATG 1333

Qy 128 CCACAACATCCGAGTGCAGGCTCGCGGAGAGAGAGGTCACTTTGACAGACTGCAAG 187
Db 1334 CCACAACATCTCGAGCGCAGCGCTCGGCGAGAGAGGTCACTTTGACAGACTGCAAG 1393

Qy 188 TCCTGGACGATCATTACCGGACGTGCTCAAGGAGATGAAGGCGTAGGCGTCCACAGTGA 247
Db 1394 TCCTGGACGACCACTACCGGACGTGCTCAAGGAGATGAAGGCGTAGGCGTCCACAGTGA 1453

Qy 248 AGGCTAAACTGCTATCTGTAGAGAGCATGCAAGCTGACGCCCCCGCAATTTCGGCCAAAT 307

1454 AGGCTAAACTCTCTATCCGTAGAGGAAGCCTGCAAGCTGACGCCCCACATTCGGCCAAAT 1513
Qy 308 CCAAAATTTGGCTATGGGCAAAAGGACGTCGGAGCCTATCCAGCAGGCGGTAAACCACA 367
Db 1514 CCAAGTTTGGCTATGGGCAAAAGGACGTCGGAGCCTATCCAGCAGGCGGTAAACCACA 1573
Qy 368 TCCGCTCCGTGTGGAAGGACTTGTCTGGAGGACACTGCACACACCAATTCAGACACCATCA 427
Db 1574 TCCACTCCGTGTGGAAGGACTTGTCTGGAAGACACTGTGACACCAATTCAGACACCATCA 1633
Qy 428 TGGCAAAAAATAGGTTTCTGCGTCAACACAGAGAAAGGAGCGCCCAACACAGCTCGCC 487
Db 1634 TGGCAAAAAATGAGTTTCTGTGTCCACACAGAGAAAGGAGCGCCGTAAAGCCAGCCGCC 1693
Qy 488 TCATCGTATTTCCAGACCTGGGAGTTCTGTGTATGCCAGAAAGATGGCCCTTTACGACGTGG 547
Db 1694 TTATCGTATTTCCAGATCTGGGAGTCGGTGTATGCCAGAAAGATGGCCCTCTATGATGTGG 1753
Qy 548 TTTCCACTCTTCTCAGGCGGTGATGGGCTCTCTCATACGGATTCGAATCTCTCTTAAGC 607
Db 1754 TCTCCACCTTCTCAGTCTGATGGGCTCTCTATACGGATTCAGTACTCTCTCTGGGC 1813
Qy 608 AGCGGTCGAGTTTCTGCTGAATACCTGGAAGCAAGAAATGCGCTATGGGCTTCTCAT 667
Db 1814 AGCGAGTCGAGTTTCTGCTGAATACCTGGAATTCNAAGAAAAACCCCATGGGCTTTTCAT 1873
Qy 668 ATGACACCCGCTGTTTGTACTCAACGCTCACTGAGATGACATCCGTTTGAGGAGTCAA 727
Db 1874 ATGACACTCGCTGTTTGTACTCAACGCTCACCGAGAACGACATCCGTTTGAGGAGTCAA 1933
Qy 728 TTTTACCAATGTTGTGACTTGCSCCCCGAAGCTAGACAGGCGCATAAAGTCTCGTCACAGAGC 787
Db 1934 TTTTACCAATGTTGTGACTTGTGCCCCCGAAGCCAGACAGGCGCATAAATCGTTCACAGAGC 1993
Qy 788 GGCTCTATGTCCGGGTCCTCATGACTAACTCCAAAGGGCGAGAACTGCGGCTATCGCCGT 847
Db 1994 GCTTTTATATCGGGGTCCTCTGACTAATTCAAAGGGCGAGAACCTGCGGTTATCGCCGT 2053
Qy 848 GCCGCGAGCGCGCTGTGACGACTAGCTGCGGTAATACCTCACATGCTACTTTGAAGG 907
Db 2054 GCCGCGAGCGCGCTGTGACGACTAGCTGCGGTAACACCTCACATGTTTACTTTGAAGG 211
Qy 908 CCGCTGACGCTGTGAGCTGCCAAGCTCCAGACTCCAGACTGCACGATGCTGCTGAATGAGAGC 967
Db 2114 CTTCTGACGCTGTGAGCTTCCGAAGCTCCAGACTCCAGACTGCACGATGCTGCTGAATGAGAGC 2177
Qy 968 ACCTTGTGCTTATCTGTGAAAGCGGGAGACCCCAAGAGGAGCGCGCAAGCCTACGAGTCT 102
Db 2174 ACCTGCTGCTTATCTGTGAAAGCGGGAGAACCCCAAGAGGAGCGCGGAGACCTACGAGTCT 223
Qy 1028 TCACGAGGCTATGACTAGTACTCTGCCCCCTTGGGGAGCCCGCCCCCAACCGGAATACG 108
Db 2234 TCACGAGGCTATGACTAGTACTCTGCCCCCTTGGGGAGCCCGCCCCCAACCGGAATACG 229
Qy 1088 ACTTGGAGCTGATAACATCGTGTTCCTCAATGTGTCGTGGCACACCATGCATCTGSCA 114
Db 2294 ACTTGGAGCTGATAACATCATGTTCTCTCAATGTGTCGTGGCCACCATGCATCAGGCA 235
Qy 1148 AAAGGGTGTACTACTTACCTACCGGTGACCCCGTCCCGTTCGGCGGGGTGGTGGGAGA 120
Db 2354 AAAGGGTGTACTACTTACCTACCGGTGATCCCGTTCGGGTGATCCCGTTCGGGTGGGAGA 241
Qy 1208 CAGCTAGGCACACTCCAGTCAACTCTGCTAGGCAACATCATCATATGATGTCGCCCACTT 126
Db 2414 CAGCTAGACACACTCAAGTTAACTCTGGCTAGGCAACATTAATATGATGTCGCCCACTT 247
Qy 1268 TGTGGCAAGGATGATTTCTGATGACTCTTCTTCATCTCTTCCATCTCTAGCCAGGAGCAAC 132
Db 2474 TGTGGCAGGATGATTTCTGATGACTCACTTCTTCCATCTCTAGCCAGGAGCAAC 253
Qy 1328 TTGAAAAAGCCCTGGATTTGTGATCTACGGGCTTGTACTTCCATTGAGCCACTTGACC 138
Db 1328 TTGAAAAAGCCCTGGATTTGTGATCTACGGGCTTGTACTTCCATTGAGCCACTTGACC 138

Db 2534 TTGAAAAGCCCTGGAGTCCAGATCTACGGGCGCTGTACTCCATTGAGCCACTTGACC 2593
Qy 1388 TACCTCAGATCATTTGAACGACTCCATGGTCTTAGCGCAATTTTCACTCCATAGTACTCTC 1447
Db 2594 TACCTCAGATCATTTGAACGACTCCATGGCCTTAGCGCAATTTTCACTCCATAGTACTCTC 2653
Qy 1448 CAGGTGAGATCAATAGGTTGGCTTCACTGCTCAGGAAGCTTGGGGTACACCCCTTGCGAG 1507
Db 2654 CAGGTGAGATCAATAGGTTGGCTTCACTGCTCAGGAAGCTTGGGGTACACCCCTTGCGAG 2713
Qy 1508 TCTGGAGACATCGGGCCAGAGTGTCCGGCTTAAGTTACTGTCCAGGGGGGAGGGCGG 1567
Db 2714 TCTGGAGACATCGGGCCAGAGAGCTCCGGCTAGTCTGTCCAGGGAGGGAGGGCGG 2773
Qy 1568 CCAATTGTGGCAAGTACCTCTTTCAACTGGGCAAGTAAGCAACCAAGCTTAAACTCACTCCAA 1627
Db 2774 CCAATTGTGGCAAGTACCTCTTTCAACTGGGCAAGTAAGCAACCAAGCTTAAACTCACTCCAA 2833
Qy 1628 TTCCGGCTGCGTCCGGCTGGACTGTTCGGGCTGGTTTGGTGGCTGCTACAGCGGGGAG 1687
Db 2834 TCCCGGCTGCGTCCGGGCTGGACTGTTCGGGCTGGTTTGGTGGCTGCTACAGCGGGGAG 2893
Qy 1688 ACATATATCACAGCCTGTCTGTCGCCGACCCCG 1721
Db 2894 ACATATATCACAGCCTGTCTGTCGCCGACCCCG 2927

RESULT 4

US-08-384-616-49
; Sequence 49, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.

Query Match 88.5%; Score 1539.6; DB 2; Length 2991;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1605; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 8 GCATGTCAATGTCTTATACATGACAGCGCGCTGTATCACACCGTGGCTGCGGAGGAAA 67
Db 1214 GCTGCTCAATGTCTTACACATGACAGCGCGCTGTATCACGCGCATGCGTGGGAGGAAA 1273
Qy 68 GCAAGCTGCCCATCAACCGCTTGAGCAACTCTTGTGCTGCTACCATTAACCTGTCTATT 127
Db 1274 GCAAGCTGCCCATCAACCGCTTGAGCAACTCTTGTGCTGCTACCATTAACCTGTCTATT 1333
Qy 128 CCACACATCCCGCAGTGCAGCTTGGCGCAGAGAAGGTCACTTTGACAGACTGCAAG 187
Db 1334 CCACACATCCCGCAGCTTGGCGCAGAGAAGGTCACTTTGACAGACTGCAAG 1393
Qy 188 TCCTGGACACATCAATACCGGGAGCTGCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTGA 247
Db 1394 TCCTGGACACATCAATACCGGGAGCTGCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTGA 1453
Qy 248 AGGCTAAACTGCTATCTGTAGAAAGAGCATGCAAGCTGACCGCCCGCATTTCCGCCAAAT 307
Db 1454 AGGCTAAACTGCTATCCGTAGAGAGAGCTGCAAGCTGACCGCCCGCATTTCCGCCAAAT 1513
Qy 308 CCAAAATTTGGCTATGGGGCAAGAGAGCTGCGGAGCTTATCCAGCAGGCGCGTTAACACA 367
Db 1514 CCAAGTTTGGCTATGGGGCAAGAGAGCTGCGGAGCTTATCCAGCAGGCGCGTTAACACA 1573
Qy 368 TCCGCTCCGTGTGGAAGGACTTGTCTGGAGAGACTGACACACCAANTTACAGACCACTCA 427
Db 1574 TCCACTCCGTGTGGAAGGACTTGTCTGGAGAGACTGACACCAANTTACAGACCACTCA 1633
Qy 428 TGGCAAAAATAGAGTTTCTGCTGCCAACAGAGAAAGAGCGCCGCAACAGCTCGCC 487
Db 1634 TGGCAAAAATAGAGTTTCTGCTGCCAACAGAGAAAGAGCGCCGCAACAGCTCGCC 1693
Qy 488 TCATCGTATTTCCAGACCTGGGAGTTTGTGTATGCGAGAAGATGGCGCTTTAGAGCTGG 547
Db 1694 TTATCGTATTTCCAGATCTGGGAGTTCGGTGTATGCGAGAAGATGGCGCTTTATGATGG 1753
Qy 548 TTTCCACTTCTCTAGGGCGGTGATGGGCTTCTCTATACGGATTCCAAATCTCTCTTAAGC 607
Db 1754 TCTCCACCTTCTCTAGGGCTGATGGGCTTCTCTATACGGATTCCAAATCTCTCTTAAGC 1813
Qy 608 AGCGGCTCGAGTTCTGTGTGAATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCAT 667
Db 1814 AGCGGCTCGAGTTCTGTGTGAATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCAT 1873
Qy 668 ATGACACACCGCTGTTTGTGACTCAACCGTCACTGAGAAATGACATCCGCTGTGGAGGCTCAA 727

Db 1874 ATGACACTCGCTGTTTCGACTCAACGGTCACCGAGACGACATCCGGTGTGAGGAGTCAA 1933
Qy 728 TTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCCATAAGTGCTCCTCACAGAC 787
Db 1934 TTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCCATAAAATCGCTCACAGAC 1993
Qy 788 GGCTCTATGCGGGGTCCCATGACTTAATCCAAAGGCGAGACTCGGCTATCGCGGT 847
Db 1994 GGCTTTATCGGGGTCTCTGACTTAATTCAAAAGGGCAGAACTCGGTTATCGCGGT 2053
Qy 848 GCCGCGAGCGGGGTGTGACGACTAGCTGCGTAAATACCTCACATGCTACTTGAAG 907
Db 2054 GCCGCGAGCGGGGTGTGACGACTAGCTGCGGTAAACCTCACATGCTACTTGAAG 2113
Qy 908 CCGTGTGAGCGCTGTGAGCTGCCAAGCTCCAGGACTGCAGGATGCTCGTGAATGGAGACG 967
Db 2114 CCTGTGAGCGCTGTGAGCTGCCAAGCTCCAGGACTGCAGGATGCTCGTGAACGAGACG 2173
Qy 968 ACCTGTGCTTATCTGTGAAGCGGGGAACCCAAAGAGAGCGGGCAAGCCTACGAGTCT 1027
Db 2174 ACCTGTGCTTATCTGTGAAGCGGGGAACCCAAAGAGAGCGGGCGAGCCTACGAGTCT 2233
Qy 1028 TCAGGAGGCTATGACTAGTACTCTGCCCGCCCTGGGGACCCGCCCAACCGGAATACG 1087
Db 2234 TCAGGAGGCTATGACTAGTACTCTGCCCGCCCTGGGGACCCGCCCAACCGGAATACG 2293
Qy 1088 ACTTGGAGCTGATAACATCTGTGTTCTTCCATGTGTGGTGGCACACGATGCTCGCA 1147
Db 2294 ACTTGGAGCTGATAACATCTGTGTTCTTCCATGTGTGGTGGCACACGATGCTCGCA 2353
Qy 1148 AAAGGCTTACTACTCTACCGGTGACCCCGTCCCGTCCCGTGGCGGCTGCGTGGAGA 1207
Db 2354 AAAGGCTTACTACTCTACCGGTGACCCCGTCCCGTGGCGGCTGCGTGGAGA 2413
Qy 1208 CAGCTAGGACACTCCAGTCAACTCTCGGTAGGCAACATCATGTATGCGGCCACTT 1267
Db 2414 CAGCTAGGACACTCCAGTCAACTCTCGGTAGGCAACATCATGTATGCGGCCACTT 2473
Qy 1268 TGTGGGCAAGGATGATCTGATGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1327
Db 2474 TGTGGGCAAGGATGATCTGATGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2533
Qy 1328 TTGAAAAGCCCTGATGTAGATCTACGGGCTTGTACTCCATTCAGCCACTTGACC 1387
Db 2534 TTGAAAAGCCCTGATGTAGATCTACGGGCTTGTACTCCATTCAGCCACTTGACC 2593
Qy 1388 TACCTCAGATCATGAACGACTCCATGTTTACGGCATTTTACCTCATAGTTACTCTC 1447
Db 2594 TACCTCAGATCATGAACGACTCCATGTTTACGGCATTTTACCTCATAGTTACTCTC 2653
Qy 1448 CAGGTGAGATCAATAGGTGGTTCATGCTCAGGAAGCTTGGGTATCACCCCTTGGCAG 1507
Db 2654 CAGGTGAGATCAATAGGTGGTTCATGCTCAGGAAGCTTGGGTATCACCCCTTGGCAG 2713
Qy 1508 TCTGGAGACATCGGGCGAGAGTGTCCGGCTAAGTTACTGTCCAGGGGGGAGGGCG 1567
Db 2714 TCTGGAGACATCGGGCGAGAGTGTCCGGCTAAGTTACTGTCCAGGGGGGAGGGCG 2773
Qy 1568 CCATTGTGGCAAGTACCTCTTCACTGGGCAAGTAAAGCAACCAAGCTTAACTCACTCAA 1627
Db 2774 CCATTGTGGCAATACCTCTTCACTGGGCAAGTAAAGCAACCAAGCTTAACTCACTCAA 2833
Qy 1628 TTCCGGCTCGCTCCGGCTGGACTTGTCCGGTGGTTCGTTGCTGCTACAGCGGGGAG 1687
Db 2834 TTCCGGCTCGCTCCGGCTGGACTTGTCCGGTGGTTCGTTGCTGCTACAGCGGGGAG 2893
Qy 1688 ACATATATACAGCCTGTCTCGTGGCCGACCCCG 1721
Db 2894 ACATATATACAGCCTGTCTCGTGGCCGACCCCG 2927

; Sequence 49, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
; ADDRESS: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeLland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: 1..2991 /note: "sequence = 6372 - 9362 of
; OTHER INFORMATION: SEQ ID NO: 1"
; NAME/KEY: CDS
; LOCATION: 1..2991
; US-08-904-686A-49

CURRENT APPLICATION DATA:		
APPLICATION NUMBER:	US/09/315,850	
FILING DATE:		
PRIOR APPLICATION DATA:		
APPLICATION NUMBER:	US/08/904,686	
FILING DATE:	01-AUG-1997	
APPLICATION NUMBER:	US/08/324,977	
FILING DATE:	18-OCT-1994	
PRIOR APPLICATION DATA:		
APPLICATION NUMBER:	JP 2-167466	
FILING DATE:	25-JUN-1990	
PRIOR APPLICATION DATA:		
APPLICATION NUMBER:	JP 2-230921	
FILING DATE:	31-AUG-1990	
PRIOR APPLICATION DATA:		
APPLICATION NUMBER:	JP 2-305605	
FILING DATE:	09-NOV-1990	
PRIOR APPLICATION DATA:		
APPLICATION NUMBER:	US 08/099,706	
FILING DATE:	30-JUL-1993	
PRIOR APPLICATION DATA:		
APPLICATION NUMBER:	US 07/769,996	
FILING DATE:	02-OCT-1991	
PRIOR APPLICATION DATA:		
APPLICATION NUMBER:	US 07/635,451	
FILING DATE:	28-DEC-1990	
ATTORNEY/AGENT INFORMATION:		
NAME:	McLeland, Le-Nhung	
REGISTRATION NUMBER:	31,541	
REFERENCE/DOCKET NUMBER:	900703G	
TELECOMMUNICATION INFORMATION:		
TELEPHONE:	(202) 659-2930	
TELEFAX:	(202) 887-0357	
INFORMATION FOR SEQ ID NO:	49:	
SEQUENCE CHARACTERISTICS:		
LENGTH:	2991 base pairs	
TYPE:	nucleic acid	
STRANDEDNESS:	single	
TOPOLOGY:	linear	
MOLECULE TYPE:	cDNA from genomic RNA	
FEATURE:		
NAME/KEY:	misc_feature	
LOCATION:	1..2991	
OTHER INFORMATION:	/note: "sequence = 6372 - 9362 of	
OTHER INFORMATION:	SEQ ID NO: 1"	
FEATURE:		
NAME/KEY:	CDS	
LOCATION:	1..2991	
US-09-315-850-49		
Query Match		88.5%; Score 1539.6; DB 3; Length 2991;
Best Local Similarity		93.8%; Pred. No. 0; Mismatches 109; Indels 0; Gaps 0;
Matches 1605; Conservative		
QY	8	GCATGTC CAATGTCCTATACATGGACAGGCGCCCTGATCATCACCGTGGCTGCGGTGGGAGGAAA 67
Db	1214	GCTGCTCAATGTCCTACACATGACAGAGGCGCCTGATCATGCGCATGCGTGGGAGGAAA 1273
QY	68	GCAAGCTGCCCATCAACGGCGCTGAGCAACTCTCTGTGCGTCAACATAACCTGGTCTATT 127
Db	1274	GCAAGCTGCCCATCAACGGGTTTGAGCAACTCTTTGTGCGCACCAACATAACATGTTTATG 1333
QY	128	CCACAACATCCCGAGTGC AAGCCTCGCGCAGAGAAAGGTCACTTTGACAGACTGCCAAG 187
Db	1334	CCACAACATCTCGCAGCGCAGGCGCTCGGCAGAGAAAGGTCACTTTGACAGACTGCCAAG 1393
QY	188	TCTGTGACGATCATTTACGGGACGTGCTCAAGGAGATGAAGGGAAGGCGTCCACAGTGA 247
Db	1394	TCCTGGACGACCACTACCGGGACGTGCTCAAGGAGATGAAGGGAAGGCGTCCACAGTTA 1453
QY	248	AGGCTAACTGCTATCTGTAGAGAAGCATGCAAGCTGACGCCCGCCGCAATTGGCCAAAT 307
Db	1454	AGGCTAACTCTCTATCCGTAGAGGAGCCTGCAAGCTGACGCCCGCCACATTTGGCCAAAT 1513
QY	308	CCAAATTTGGCTATATGGGCAAAAGGACGTC CGAGGCTATCCAGCAGGCGCGTTAAACCACA 367
Db	1514	CCAAATTTGGCTATATGGGCAAAAGGACGTC CGGAAACCTATCCAGCAGGCGCGTTAAACCACA 1573
QY	368	TCCGCTCCGTGTGGAAGGACTT GCTGGAGGACACTGACACACCAATTTAGACACCACCATCA 427
Db	1574	TCCACTCCGTGTGGAAGGACTT GCTGGAAGACACTGTGACACCAATTTGACACCACCATCA 1633
QY	428	TGGCAAAAAATGAGGTTTCT GCGTCCAAACACAGAGAAAGGAGGCCCAAAACCAAGCTCGCC 487
Db	1634	TGGCAAAAAATGAGGTTTCT GTGTCCAAACACAGAGAAAGGAGGCCCTTAAGCCAGCCGCC 1693
QY	488	TCATCGTATTTCCAGACCTGGGAGTT CGTGTATGCGAGAAGATGGCCCTTTAGCAGCTGG 547
Db	1694	TTATCGTATTTCCAGATCTGGGAGT CGGTGTATGCGAGAAGATGGCCCTCTATGATGTGG 1753
QY	548	TTTCCACTCTTCTCAGGCGCTGATGGGT CTCTCATACGGATTCGAATACTCTCTCTAAGC 607
Db	1754	TC CCACCCCTTCTCAGGTCGTATGGGCT CTCTCATACGGATTTCCAGTACTCTCTCTGGGC 1813
QY	608	AGCGGTTCGAGTTCTTGCTCAATACCTGGAAGCAAGAAATGSCCCTATGGGCTTCTCAT 667
Db	1814	AGCGAGTCGAGTTCTTGCTGAATACCTGGAAATCAAAAGAAAAACCCCATGGGCTTTTCAT 1873
QY	668	ATGACACCCGCTGTTT GACTCAACCGTCACTCAGAAATGACATCCGCTGTGTGAGGAGTCAA 727
Db	1874	ATGACACTCGCTGTTT CGACTCAACCGTCAACCGTCAACCGAAGACGACATCCGTTGTGAGGAGTCAA 1933
QY	728	TTTACC AAATGTTGTGACTTTGGCCCCGGAAGCTAGACAGGCGCATTAAGTCTGCTCACAGAGC 787
Db	1934	TTTACC AAATGTTGTGACTTTGCCCCCCGAAGCCAGACAGAGCCATAAAATCGCTCACAGAGC 1993
QY	788	GGCTCTATGTCGGGGTCCCATGACT MACTCCAAAGGGGAGAGACTCGGGCTATTCGCGGT 847
Db	1994	GGCTTTATATCGGGGGTCTCTGACT AATTTCAAAGGGGCGAGAACTCGCGTTATTCGCGCGGT 2053
QY	848	GCCGCGGAGGGGGTGCTGACGACT AGCTGCGGTATATACCTTCACATGCTACTTTGAAGG 907
Db	2054	GCCGCGGAGGGGGTGCTGACGACTAGCT GCGGTAAACCCCTTCACATGCTACTTTGAAGG 2113
QY	908	CCGCTGCAGCCTGTCGAGCTGCCAAGCT CCAGGACTGACGATGCTCTCGTAATGGAGAGC 967
Db	2114	CCTCTGAGCCTGTCGAGCTGCGAAGCT CAGGACTGACGATGCTCGTGAACGAGAGC 2173
QY	968	ACCTTGTCTTATCTGTGAAAGCGCGGAACCC AAGAGGACGCGGCAAGCCTTAGAGTCT 1027
Db	2174	ACCTCGTCGTTTATCTGTGAAAGCGCGGAACCC AAGAGGACGCGGCGAGCCTTAGAGTCT 223
QY	1028	TCACGAGGCTATGACTAGGTACT CTGCCCCCTTGGGGACCCCGCCCAACCGGAATACG 108
Db	2234	TCACGAGGCTATGACTAGGTACT CCGCCCCCTTGGGGACCCCGCCCAACCGGAATACG 229
QY	1088	ACTTGGAGCTCATTAACATCGTGTTCCT CCATATGTCGTCGACACGATGCATGCATCTGCA 1147
Db	2294	ACTTGGAGCTCATTAACATCATATGTC CTCCAAATGTCGTCGACACGATGCATGCATCTGCA 235
QY	1148	AAAGGGTGTACTACTCTACCCCGTCACCC CACCGCTCCCTTTCGCGGGCTGCGTGGGAGA 120
Db	2354	AAAGGGTGTACTACTCTACCCCGTGTAT CCACACCCCTTAGCAGCGGCTGCGTGGGAGA 241
QY	1208	CAGCTTAGGCACACTCCAGTCAACT CTGGGCTAGGCAACATCATCATGTATGGGCCACTT 126
Db	2414	CAGCTTAGACACACTCCAGTTAACT CTGGCTAGGCAACATTAATATGTATGCGGCCACTT 247
QY	1268	TGTGGGCAAGGATGATTTCTGATGACT CACTTCTTCTCATCTCTTAGCCAGGAGCAAC 132
Db	2474	TGTGGGCAAGGATGATTTCTGATGACT CACTTCTTCTCATCTCTTAGCCAGGAGCAAC 253
QY	1328	TTGAA AAGCCCTCGATTGTGAGATCT ACGGGGCTTGTACTCCATTGAGGCACCTTGACC 138
Db	2534	TTGAA AAGCCCTTGACACTGCCAGAT CTTACGGGGCTGTTACTCCATTGAGGCACCTTGACC 259

QY 1388 TACCTCAGATCATTTGAACGACTCCATGCTTAGCGCATTTTCACTTCATCTACTCTC 1447
Db |||||
QY 2594 TACCTCAGATCATTTGAACGACTCCATGCTTAGCGCATTTTCACTTCATCTACTCTC 2653
Db |||||
QY 1448 CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAAGCTTGGGTACCAACCTTTGGAG 1507
Db |||||
QY 2654 CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAAGCTTGGGTACCAACCTTTGGAG 2713
Db |||||
QY 1508 TCTGGAGACATCGGCCAGAACTGTCGGCTTAAGTTACTGTGCCAGGGGGAGGGCCG 1567
Db |||||
QY 2714 TCTGGAGACATCGGCCAGAGAGCTGTCGGCTTAGCTAGCTAGCTAGCTAGCTAGCT 2773
Db |||||
QY 1568 CCAATTGTGGCAAGTACCTTTCAACTGGGCACTGTCGGCTTAGCTAGCTAGCTAGCT 1627
Db |||||
QY 2774 CCAATTGTGGCAAGTACCTTTCAACTGGGCACTGTCGGCTTAGCTAGCTAGCTAGCT 2833
Db |||||
QY 1628 TTCCGGCTGGCTCCCGCTGGACTGTCGGCTTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 1687
Db |||||
QY 2834 TCCCGGCTGGCTCCCGCTGGACTGTCGGCTTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 2893
Db |||||
QY 1688 ACATATATCAGACCTGTCGTCGTCGCCAGCCCG 1721
Db |||||
QY 2894 ACATATATCAGACCTGTCGTCGTCGCCAGCCCG 2927
Db |||||

RESULT 7

US-08-324-977-35
; Sequence 35, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Ise
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7863 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7863
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..7863
; OTHER INFORMATION: /note= "sequence = 1500 - 9362 of
; OTHER INFORMATION: SEQ ID NO: 1"
US-08-324-977-35
Query Match 88.5%; Score 1539.6; DB 1; Length 7863;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1605; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 8 GCATGTAATGTCTATACATGACAGGCGCCCTGATCACCGTGGCTGGCGAGGAAA 67
Db |||||
QY 6086 GCTGCTCAATGCTCTACATGACAGGCGCTTGATCAGCCATGGCTGGCGAGGAAA 6145
Db |||||
QY 68 GCAAGTGGCCATCAACGGCTGAGCAACTCTTGTGCTGCTACCATTAACCTGCTATT 127
Db |||||
QY 6146 GCAAGTGGCCATCAACGGCTGAGCAACTCTTGTGCTGCTACCATTAACATGTTATG 6205
Db |||||
QY 128 CCACACATCCCGCAGTGAAGCTCGGCGAGAGAGGTCACTTTGACAGACTGCAAG 187
Db |||||
QY 6206 CCACACATCTCGCAGCGCAGGCTCGGCGAGAGAGGTCACTTTGACAGACTGCAAG 6265
Db |||||
QY 188 TCCTGACGATCATTTACCGGCGCTCTCAAGGAGATGAAGCGAGAGGCTCCACATGA 247
Db |||||
QY 6266 TCCTGACGACCATTAACCGGCGCTCTCAAGGAGATGAAGCGAGAGGCTCCACATGA 6325
Db |||||
QY 248 AGCTAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCGCCCGCATTCGCCCAAT 307
Db |||||
QY 6326 AGCTAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCGCCCGCATTCGCCCAAT 6385
Db |||||
QY 308 CCAATTTGGCTATGGGCGAAAGAGCTCGGAGGCTATCCAGCAGGGCGTTAACCA 367
Db |||||
QY 6386 CCAAGTTTGGCTATGGGCGAAAGAGCTCGGAGACCTATCCAGCAGGGCGTTAACCA 6445
Db |||||
QY 368 TCCTGCTCGGTGGAGGACTTCTGCTGAGGAGCACTGACACACCAATTCAGACCATCA 427
Db |||||
QY 6446 TCCTGCTCGGTGGAGGACTTCTGCTGAGGAGCACTGACACACCAATTCAGACCATCA 6505
Db |||||
QY 428 TGGCAAAAAATGAGGTTTCTGCTGCAACAGAGAGAGGCGCGCAACAGCTCGCC 487
Db |||||
QY 6506 TGGCAAAAAATGAGGTTTCTGCTGCAACAGAGAGAGGCGCGTAAGCGCGCGCC 6565
Db |||||
QY 488 TCATCGTATTTCCAGACCTGGGAGTTCGTGTATGCGAGAGATGGCCCTTTACGACGTGG 547
Db |||||
QY 548 TTTCCACTTCTCCTCAGGCGGTGATGGGCTCTCATACCGATTTCGAATCTCTCTCAAGC 607
Db |||||
QY 6626 TCTCCACCTTCTCCTCAGGCTGGGCTCTCTCATACCGATTTCGAGTACTCTCTGGGC 6685
Db |||||
QY 608 AGCGGCTCGAGTTCCTGCTGATACCTGGAAGCAAAATGCTTATGGGCTTCTCAT 667
Db |||||
QY 6686 AGCGAGTCGAGTTCCTGCTGATACCTGGAAGCAAAATGCTTATGGGCTTCTCAT 6745
Db |||||
QY 668 ATGACACCCGCTGTTTGTGACTCAACGGTCACTGAGATGACATCGGCTTGAGGAGTCA 727
Db |||||

6746 ATGACACTCGTGTTCGACTCAACGGTCACCGAAGACGACATCCGTTGAGGAGTCAA 6805
728 TTTACCAATGTTGACTTGGCCCCGAGAGCTAGACAGGCCATAAGTCTCGCTACAGAGC 787
6806 TTTACCAATGTTGACTTGGCCCCGAGAGCTAGACAGGCCATAAAATCGCTACAGAGC 6865
788 GGTCTATGTCGGGGTCCCATGACTTACCTCAAGGCGCAAGCTCGGGCTATCGCCGGT 847
6866 GCTTTATATCGGGGTCTCTGACTAATTTAAAAGGGCAGAACTGGGGTTATCGCCGGT 6925
848 GCCGGCGAGCGGGTGTGACGACTAGCTCGCGTAAATACCTCACATGCTTACTTGAAGG 907
6926 GCCGGCGAGCGGGTGTGACGACTAGCTCGGGTAACACCTCACATGTTACTTGAAGG 6985
908 CCGCTGCAGCCCTGTCGAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAATGAGAGC 967
6986 CCTCTGCAGCCCTGTCGAGCTCGGAGCTCCAGGACTCGCAAGTCTCGTGAACGAGAGC 7045
968 ACCTTGTCTTATCTGTGAAGCGGGGAAACCAAGAGGACGGCGCAAGCCTACGAGTCT 1027
7046 ACCTCGTCTTATCTGTGAAGCGGGGAAACCAAGAGGACGGCGCAAGCCTACGAGTCT 7105
1028 TCACGGAGCTATGACTAGTACTTCTGCCCGCTCGGGAGCCGCCCAACCGGAATACG 1087
7106 TCACGGAGCTATGACTAGTACTTCTGCCCGCTCGGGAGCCGCCCAACCGGAATACG 7165
1088 ACTTCGAGCTGATAACATGTTCTTCTCAATGTTGCTCGGTGGCAGACCATGATCTGGCA 1147
7166 ACTTCGAGCTGATAACATGTTCTTCTCAATGTTGCTCGGTGGCAGACCATGATCTGGCA 7225
1148 AAAGGGTGTTACTACCTCACCCGTGACCCACCGTCCCTTGGCGGGCTGGTGGGAGA 1207
7226 AAAGGGTGTTACTACCTCACCCGTGATCCCAACCCCTAGCAGGGCTGGTGGGAGA 7285
1208 CAGCTAGGACACTCCAGTCAACTCTGCTGCTAGGCAACATCATGATGATGGCCCACTT 1267
7286 CAGCTAGACACACTCCAGTCAACTCTGCTGCTAGGCAACATTAATGATGCGCCCACTT 7345
1268 TGTGGGCAAGGATGTTCTGATGACTCACTTCTTCTCCATCTCTTAGCCAGGAGCAAC 1327
7346 TGTGGGCAAGGATGTTCTGATGACTCACTTCTTCTCCATCTCTTAGCCAGGAGCAAC 7405
1328 TTGAAAAGCCCTGGATTTCTAGATCTACGGGGCTTGTACTCCATTTAGCCCACTTGACC 1387
7406 TTGAAAAGCCCTGGATTTCTAGATCTACGGGGCTTGTACTCCATTTAGCCCACTTGACC 7465
1388 TACCTCAGATCATTTGAACGACTCCATGGTCTTAGGCGCAATTTTCACTCCATGATGTTACTCTC 1447
7466 TACCTCAGATCATTTGAACGACTCCATGGTCTTAGGCGCAATTTTCACTCCATGATGTTACTCTC 7525
1448 CAGGTGAGATCAATAGGGTGGCTTCTGCTCAGGAGCTTGGGGTACCACTTGGCGAG 1507
7526 CAGGTGAGATCAATAGGGTGGCTTCTGCTCAGGAGCTTGGGGTACCACTTGGCGAG 7585
1508 TCTGGAGACATCGGGCCAGAGTGTCCGGCTAAGTTACTGTCCAGGGGGAGGGCGG 1567
7586 TCTGGAGACATCGGGCCAGAGGCTTCCGGCTAGGCTTACTGTCCAGGGAGGAGGGCGG 7645
1568 CCATTTGTGCAAGTACCTCTTCAACTGGGCACTAAGGACCAAGCTTAAACTCACTCCAA 1627
7646 CCATTTGTGCAAGTACCTCTTCAACTGGGCACTAAGGACCAAGCTTAAACTCACTCCAA 7705
1628 TCCGGCTGGCTCCCGCTCGGACTTGTCCGGCTGGTGTGCTGCTGCTAGCGGGGAG 1687
7706 TCCGGCTGGCTCCCGCTCGGACTTGTCCGGCTGGTGTGCTGCTGCTAGCGGGGAG 7765
1688 ACATATATCAGGCTGTCTGCTGCCGACCCCG 1721
7766 ACATATATCAGGCTGTCTGCTGCCGACCCCG 7799

US-08-384-616-35
Sequence 35, Application US/08384616
Patent No. 5847101
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
ADDRESS: Naughton
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,616
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 7863 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7863
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..7863
OTHER INFORMATION: /note= "sequence = 1500 - 9362 of
OTHER INFORMATION: SEQ ID NO: 1"
US-08-384-616-35

Query Match 88.5%; Score 1539.6; DB 2; Length 7863;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1605; Conservative 0; Mismatches 109; Indels 0; Gaps 0
9Y 8 GCATGTCAATGTCCTATATACATGACAGCGCCCTCATCACACCGTCCGCTGCGGAGGAAA 67

	Db	6086	GCTGCTCAATGTCCTTACACATGGACXAGGCGCCTTGATCACGCCAATGCGCTGCGAGGAAA	6145
	Qy	68	GCAAGCTGCCCATCAAACGCGCTGAGCAAACTCTCTGCTGCCTGCAACCATAAACCTGTGCTATT	127
	Db	6146	GCAAGCTGCCCATCAAACGCGTTGAGCAACTCTTTGCTGCGCCACCAATAACAATGTTTTATG	6205
	Qy	128	CCACAACATCCCGCGAGTGAAGCCTTCGGCGCAAGAAGGTACCTTTGACAGATGCAAG	187
	Db	6206	CCACAACATCTCGCAGCGCAGCGCTTCGGCGAAGAGGTCACTTTGACAGATGCAAG	6265
	Qy	188	TCTCTGACGATCAATTAACCGGAGCTGCTCAAGGAGATGAAGGCGAAGCGCTCCAAGTGA	247
	Db	6266	TCTCTGACGACCACTTAACCGGAGCTGCTCAAGGAGATGAAGGCGAAGCGCTCCACAOTTA	6325
	Qy	248	AGGCTAAACTGCTATCTGTAGAAGAGCATCAAGCTGACGCCGCCCGCATTCGGCCMAAT	307
	Db	6326	AGGCTAAACTCTATCTCGTAGAGGAAGCTCAAGCTGAGCGCCCCCACTTCGGCCMAAT	6385
	Qy	308	CCAAATTTGGCTATTCGGGCAAGAGACGTCGGAGCCTATCCAGCAGAGCGCGTTAAACACA	367
	Db	6386	CCAACTTTGGCTATTCGGGCAAGAGACGTCGGGAACCTATCCAGCAGAGCGCGTTAAACACA	6445
	Qy	368	TCGGTTCGGTGTGAAGGACTTGTCTGGAGGACACTGACACACCAATTTACAGACCACATCA	427
	Db	6446	TCCACTCCGGTGTGAAGGACTTGTCTGGAGGACACTGTGACACCAATTTGACACACCAATCA	5505
	Qy	428	TGGCAAAAAATAGAGTTTTCGCGTCCAAACAGAGAAAGGAGGCGCAAAACCAAGCTCGCC	487
	Db	6506	TGGCAAAAAATAGAGTTTTCGCGTCCAAACAGAGAAAGGAGGCGGTAAACCAAGCGCGC	5565
	Qy	488	TCATCGTATTCACAGACCTGGAGTTTGTGTATCGAGAAAGATGAGCCCTTTACGAGCTGG	547
	Db	6566	TTATCGTATTCACAGATCTGGAGTCCGTTGTATCGAGAAAGATGGCCCTCTATGATGTGG	6625
	Qy	548	TTTCCACTCTTCTTCAGCGCGTGAATGGCTCCTCATACGATTTCCAATACTCTCTTAAGC	607
	Db	6626	TCTCCACCCTTCTTCAGGCTGTGAATGGCTCCTCATACGATTTCCAAGTACTCTCTCTGGGC	6685
	Qy	608	AGCGGTCGAGTTCCTGTGTAAATACCTTGGAAAGCAAGAAATCGCTATGGGCTTCTCAT	667
	Db	6686	AGCGAGTCGAGTYCTGGTGAATACCTTGBAAATCAAGAANAACCCCATGGGCTTTTCAT	6745
	Qy	668	ATGACACCGCTGTTTTGACTCAAACGCTCACTGAGAATGACATCCGTGTTTGAGGAGTCAA	727
	Db	6746	ATGACACTCGCTGTTTCGACTCAACGCTACCGAGAACGACATCCGTGTTTGAGGAGTCAA	6805
	Qy	728	TTTACCAATGTTGTGACTTGCGCCCCGAAAGCTAGACAGGCCATPAAGSTCGCTCACAGAGC	787
	Db	6806	TTTACCAATGTTGTGACTTGCGCCCCGAAAGCTAGACAGGCCATPAAAATCGCTCACAGAGC	6865
	Qy	788	GGCTCTATGTCGGGGTCCCATGNCTAACTCCAAAGGCGNGAACTCGCGGCTATCGCCGT	847
	Db	6866	GGCTTTTATTCGGGGGTCCCTCTGACTAAATTCAAAAGGGCGAGAACTCGCGGTATTCGCGGT	6925
	Qy	848	GCCCGCGAGCGGCGTCTGACGACTAGCTGCGGTAAATACCTTCATGCTACTTCTTAAGG	907
	Db	6926	GCCCGCGAGCGGCGTCTGACGACTAGCTGCGGTAAATACCTTCATGCTACTTCTTAAGG	6985
	Qy	908	CCGCTGACGCTGTTCGAGCTGCCAAGCTCCAGGACTTGCACGATGTCTGTAATGGAGAGC	967
	Db	6986	CCTCTGACGCTGTTCGAGCTTGGAGCTTCAGAGCTTGCAGATGTCTGHAACGGAGAGC	7045
	Qy	968	ACCTTGTGTTATCTGTGAAGCGCGGAAACCCAGAGGACGCGGCAAGCTTACGAGTCT	1027
	Db	7046	ACCTCGTGTATCTGTGAAGCGCGGAAACCCAGAGGACGCGGAGCTTACGAGTCT	7105
	Qy	1028	TCAGGAGGCTATGACTAGTACTCTGCCCCCTCTGGGAGCCGCGCCCAACCGGAATACG	1087
	Db	7106	TCAGGAGGCTATGACTAGTACTCTGCCCCCTCTGGGAGCCGCGCCCAACCGGAATACG	7165
	Qy	1088	ACTTTGAGCTGATAACATCGTGTTCCTCCAAATGTGTGCTGCCACAGATGCACTTGCA	1147

Db	7166	ACTTGGAGCTGATAACATCATCTGTTCTCTCCAATGTGTCGGTGCGCCACGATGTCATAGCCA	7225
QY	1148	AAAGGGTGTACTACTCAGCCGGTGAACCCACCGTCCCGCTTGGCGGGCTGCGTGGGAGA	1207
Db	7226	AAAGGGTGTACTACTCAGCCGGTATCCACACCCCGCTAGCACGGGCTGCGTGGGAGA	7285
QY	1208	CAGCTTAGGCACACTCCAGTCAACTCCTGGCTTAGGCAACATCATCATATATGCGGCCACTT	1267
Db	7286	CAGCTTAGACACACTCCAGTTAACTCCTGGCTAGGCAATTTATGATGCGGCCACTT	7345
QY	1268	TGTGGGCAAGATGATTTCTGATGACTCAGCTTCTTCTCATCTCTTAGGCCAGGAGCAAC	1327
Db	7346	TGTGGGCAAGATGATTTCTGATGACTCAGCTTCTTCTCATCTCTTAGGCCAGGAGCAAC	7405
QY	1328	TTGAAAAGCCCTGGATTTGTTCAGATCTACGGGCTGTTTACTCCTCATTTAGGCCACTTGACC	1387
Db	7406	TTGAAAAGCCCTGGATCTCCAGATCTACGGGCTGTTTACTCCTCATTTAGGCCACTTGACC	7465
QY	1388	TACCTCAGATCATTTGAACGATCTCATGCTTTTAGCCGATTTTTCACCTCCATAGTACTCTC	1447
Db	7466	TACCTCAGATCATTTGAACGATCTCATGCTTTAGCCGCTTAGCCGATTTTTCACCTCCATAGTACTCTC	7525
QY	1448	CAGGTGAGATCAATAGGGTGCTTTCATGCTCTAGGAAGCTTGGGGTACCACCTTGGGAG	1507
Db	7526	CAGGTGAGATCAATAGGGTGCTTTCATGCTCTAGGAAGCTTGGGGTACCACCTTGGGAG	7585
QY	1508	TCTGGAGACATCGGGCCAGAAAGTGTCCGGCTAAAGTTACTGTCCACGGGGGAGGGCCG	1567
Db	7586	TCTGGAGACATCGGGCCAGGAGCTCCGGCTAGGCTACTGTCCACGGGAGGGAGGGCCG	7645
QY	1568	CCATTTGTGGCAAGTACTCTTCAACTGGGCAGTAAGGACCAAGCTTAAACTCATCTCCAA	1627
Db	7646	CCATTTGTGGCAAAATACCTTTCAACTGGGCAGTAAAAACCAAACTTAAACTCACTCCAA	7705
QY	1628	TTCCGGCTGGTCCCGGCTGGACTTGTTCGGCTGGTTTGGTGTGCTACACGGGGGAG	1687
Db	7706	TCCCGGCTGGTCCCGGCTGGACTTGTTCGGCTGGTTTGGTGTGCTAACGGGGGAG	7765
QY	1688	ACATATATCAGACCTGTCTCTGTCGCCGACCCCG	1721
Db	7766	ACATATATCAGACCTGTCTCTGTCGCCGACCCCG	7799

```

1  RESULT 9
2  US-08-904-686A-35
3  ; Sequence 35, Application US/08904686A
4  ; Patent No. 5998130
5  ; GENERAL INFORMATION:
6  ; APPLICANT: OKAYAMA, Hiroto
7  ; APPLICANT: FUKU, Isao
8  ; APPLICANT: MORI, Chisato
9  ; APPLICANT: TAKAMIZAWA, Akahisa
10 ; APPLICANT: YOSHIDA, Iwao
11 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
12 ; TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
13 ; NUMBER OF SEQUENCES: 50
14 ; CORRESPONDENCE ADDRESS:
15 ; ADDRESSEE: Armsstrong, Westerman, Hattori, McLeiland &
16 ; ADDRESSEE: Naughton
17 ; STREET: 1725 K St. N.W. Suite 1000
18 ; CITY: Washington
19 ; STATE: D.C.
20 ; COUNTRY: U.S.A.
21 ; ZIP: 20006
22 ; COMPUTER READABLE FORM:
23 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
24 ; COMPUTER: IBM PC compatible
25 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
26 ; SOFTWARE: ASCII
27 ; CURRENT APPLICATION DATA:
28 ; APPLICATION NUMBER: US/08/904,686A
29 ; FILING DATE: 01-AUG-1997
30 ; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: McLeiland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 7863 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7863
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..7863
OTHER INFORMATION: /note= "sequence = 1500 - 9362 of
OTHER INFORMATION: SEQ ID NO: 1"
US-08-904-686A-35
Query Match 88.5%; Score 1539.6; DB 2; Length 7863;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1605; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 8 GCATGTCATGTCCTATACATGACAGGCGCCTGATACACCGTGGCGTGGGAGGAAA 67
DB 6086 GCTGCTCAATGTCCTACATGACAGGCGCCTTGTATCAGCGCATGCGCTGGGAGGAAA 6145
QY 68 GCAAGCTGCCATCAACGGCCTGAGCAACTCCTTGTCTGGGTCAACATACTGGTCTATT 127
DB 6146 GCAAGCTGCCATCAACGGCTTGAGCAACTTTTGTGCGCCACCATACATGTTTATG 6205
QY 128 CCAACAATCCCCGAGTGCAGAGCCTGCGGCAGAGAAGGTGCACCTTTGACACAGTGCAG 187
DB 6206 CCACACATCTCGCAGCGCAGGCGCTGCGGCAGAGAAGGTGCACCTTTGACAGAGTGCAG 6265
QY 188 TCCTGACAGATCATATTACCGGAGCTGCTCAAGGAGATGAGGGAAGGCGTCCACAGTGA 247
DB 6266 TCCTGACGACCACTACCGGAGCTGCTCAAGGAGATGAGGGAAGGCGTCCACAGTGA 6325
QY 248 AGGTAACTGCTATCTGTAGAGAAGCATGCAAGCTGACGCGCCCGCATTCGGCCAAAT 307
DB 6326 AGGTAACTGCTATCTGTAGAGAAGCATGCAAGCTGACGCGCCCGCATTCGGCCAAAT 6385
QY 308 CCAAAATTTGGCTATGGGGCAAGGAGCTCCGGAGCCTATCCAGCAGGCGGTTAACACA 367
DB 6386 CCAAGTTTGGCTATGGGGCAAGGAGCTCCGGAGCCTATCCAGCAGGCGGTTAACACA 6445
QY 368 TCCGCTCCGTGTGAAGGACTTCTGTGAGGAGCACTGCACACCAAAATTCAGACCACCATCA 427

6446 TCCACTCCCGTGTGGAGGACTTGTGTGAAAGACACTGTGTACACCAATTTGACACCAATCA 6505
QY 428 TGGCAAAAAATGAGGTTTTTCTGGTCCACACAGAGAAAGGCGCAACACGAGCTCGCC 487
DB 6506 TGGCAAAAAATGAGGTTTTTCTGTGTCCAAACAGAGAGAAAGGAGCCGTAAAGCAGCCGCC 6565
QY 488 TCATCGTATTCCCGAGACCTGGGAGTTCGTGTATGCGAGAAAGATGGCCCTTTACGAGCTGG 547
DB 6566 TTATCGTATTCCCGAGATCTGGGAGTTCGTGTATGCGAGAGATGGCCCTCTATGATGTGG 6625
QY 548 TTTCCACTCTTCTCAGGGCGGTGATGGGCTCTCTATACCGATTTCCAATATCTCTCTAAGC 607
DB 6626 TCTCCACCTTCTCTCAGGTCTGTGTGAGTGGGCTCTCTATACCGATTTCCAGTACTCTCTCTGGGC 6685
QY 608 AGCGGTTCGAGTTCTCTGGTGAATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCAT 667
DB 6686 AGCGAGTTCGAGTTCTCTGGTGAATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCAT 6745
QY 668 ATGACACCCGCTGTTTTGTACTCAACGCTCACTGAGAAATGACATCCGTGTGTAGGAGTCAA 727
DB 6746 ATGACACTCGCTGTTTCGACTCAACGCTCACCGAGAACGACATCCGTTGTGAGAGTCAA 6805
QY 728 TTTACCAATGTTGTACTTGGCCCCCGAAGCTAGACAGGCCATAAGSTCGCTCACAGAGC 787
DB 6806 TTTACCAATGTTGTACTTGGCCCCCGAAGCCAGACAGGCCATAAAATCGCTCACAGAGC 6865
QY 788 GGCTCTATGTCGGGGGTCCCATGACTAACTCCAAAAGGCGAGAACTGGGGCTATCGCGGT 847
DB 6866 GSCCTTATATCGGGGTCTCTGACTAAATCAAAAGGGCAGAACTGCGGTATATCGCGGT 6925
QY 848 GCGCGCGAGGCGGCTGCTGACGACTAGCTGCGGTATACCTTCACATGCTTACTTGAAGG 907
DB 6926 GCGCGCGAGGCGGCTGCTGACGACTAGCTGCGGTAAACACCTCTCACATGTTACTTGAAGG 6985
QY 908 CCGCTGCAGCTGTGAGAGTCCCAAGCTCCAGAGCTGCAGAGTCTCGTGAATGAGAGCG 967
DB 6986 CCTCTGCAGCTGTGAGAGTCCCAAGCTCCAGAGCTGCAGAGTCTCGTGAATGAGAGCG 7045
QY 968 ACCTTGTCTGTTATCTGTGAAAGCGCGGAAACCCCAAGAGAGCGCGGCAAGCCTACGAGTCT 1027
DB 7046 ACCTCTGCTGTTATCTGTGAAAGCGCGGAAACCCCAAGAGAGCGCGGCAAGCCTACGAGTCT 7105
QY 1028 TCACGAGGCTATGACTAGGTACTCTGCCCGCTCTGGGAGCCCGCCCGCAACCGAATACG 1087
DB 7106 TCACGAGGCTATGACTAGGTACTCTGCCCGCTCTGGGAGCCCGCCCGCAACCGAATACG 7165
QY 1088 ACTTGGAGCTGATTAACATCGTGTCTTCCATGTTGCTGCGACACGATGSCATCTGGCA 1147
DB 7166 ACTTGGAGCTGATAAATCATGTTCTTCCAAATGTTGCGTGGCCCAAGATGCATCAGGCA 7225
QY 1148 AAAGGTGTACTACTCATCCGTCGACCCACCGTCCCGCTTGGCGGGCTGCGTGGGAGA 1207
DB 7226 AAAGGTGTACTACTCATCCCGTGTATCCACCAACCGCTGAGCAGGCGTGGGAGA 7285
QY 1208 CAGCTTAGGCACACTCCAGTCAACTCTGSGTAGGCAACATCATCATGTATGCGGCCACTT 1267
DB 7286 CAGCTTAGGCACACTCCAGTCAACTCTGSGTAGGCAACATCATCATGTATGCGGCCACTT 7345
QY 1268 TGTGGCAAGGATGATTTCTGATGACTCATCTTCTTCTCCATCTTCTAGCCAGGAGCAAC 1327
DB 7346 TGTGGCAAGGATGATTTCTGATGACTCATCTTCTTCTCCATCTTCTAGCCAGGAGCAAC 7407
QY 1328 TTGAAAAAGCCCTGATCTCAGATCTACGGGCTTGTGTACTTCATTTAGGCACTTTGACC 138
DB 7406 TTGAAAAAGCCCTGATCTCAGATCTACGGGCTTGTGTACTTCATTTAGGCACTTTGACC 746
QY 1388 TACCTTCAGATCAATTCGAAAGCTTCCATGCTTCTAGCGCATTTTACTTCATTCATAGTACTCTC 144
DB 7466 TACCTTCAGATCAATTTGAAAGCTTCCATGCGCTTAGGCAATTTTACTTCATAGTACTCTC 752
QY 1448 CAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGGTACCACTTGGCGAG 150

QY 8 GCATGTCAATGTCCTATACATGACAGGCGCCTGTATCACACCGTGGCTGGGAGAAA 67
Db 6086 GTGTCTCAATGTCTTACATGACAGGCGCCTGTATCACGCGATGCGCTGGGAGAAA 6145
QY 68 GCAAGCTGCCATCAACGCGCTGAGCAACTCTTGTGCTGCGTCAACATTAACCTGGTCTATT 127
Db 6146 GCAAGCTGCCATCAACGCGCTGAGCAACTCTTGTGCTGCGCACCATACATGCTGGTTATG 6205
QY 128 CCACAACTCCGCGAGTCAAGCTGCGGCGAGAGAGGTCACTTTGACAGACTGGCAAG 187
Db 6206 CCACAACTCTGCGAGCGAGGCTGCGGCGAGAGAGGTCACTTTGACAGACTGGCAAG 6265
QY 188 TCTTGACGATCATTAACCGGACGCTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA 247
Db 6266 TCTTGACGATCACTAACCGGAGCTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTTA 6325
QY 248 AGGCTAACTGCTATCTGTAGAGAGCATGAAGCTGAAGCGGCGGCTTGGGCAAT 307
Db 6326 AGGCTAACTGCTATCTGTAGAGAGCATGAAGCTGAAGCGGCGGCTTGGGCAAT 6385
QY 308 CCAAAATTTGGCTATGCGGCAAGGAGCTGCGGAGCTATCCAGAGGCGGCTTAACCA 367
Db 6386 CCAAGTTTGGCTATGCGGCAAGGAGCTGCGGAGCTATCCAGAGGCGGCTTAACCA 6445
QY 368 TCCGCTCCGCTGTGGAAGGACTTGTGCGGACACTGACACACCAATTCAGACCCACATCA 427
Db 6446 TCCACTCCGCTGTGGAAGGACTTGTGGAAGACACTGTGACCAATTCAGACCCACATCA 6505
QY 428 TGGCAAAAATAGAGTTTCTGCGTCCAAACAGAGAGAGGCGGCGCAACAGCTCGCC 487
Db 6506 TGGCAAAAATAGAGTTTCTGCGTCCAAACAGAGAGAGGCGGCGTAAAGCAGCGCCGCC 6565
QY 488 TCATCGTATCCAGACCTGGAGTTGCTGTATGCGAAGATGCCCTTTACGACGTGG 547
Db 6566 TTATCGTATCCAGATCTGGAGTCCGCTGTATGCGAAGATGCCCTTTATGATGG 6625
QY 548 TTTCCACTCTTCTCAGCGCTGATGGGCTCTCATACGGATTCGAATCTCTTAAAGC 607
Db 6626 TCTCACCCTTCTCAGGCTGCTGATGGGCTCTCATACGGATTCGAATCTCTTAAAGC 6685
QY 608 AGCGGCTCGAGTTCTCGTGTGAATACCTGGAAGCAAGAAATGCCCTATGGCTTCTCAT 667
Db 6686 AGCGAGTCGAGTTCTCGTGTGAATACCTGGAATCAAAGAAATGCCCTATGGCTTCTCAT 6745
QY 668 ATGACACCGCTGTTTGAATCAACGCTCACTGAGAAATGACATCGCTGTGAGAGTCAA 727
Db 6746 ATGACACTCGCTGTTTGAATCAACGCTCAACGAGAACGATCGCTGTGAGAGTCAA 6805
QY 728 TTTACCAATGTGACTTGGCCCCCGAAGCTAGACAGGCGCATAGGTCGCTCACAGAGC 787
Db 6806 TTTACCAATGTGACTTGGCCCCCGAAGCTAGACAGGCGCATAGGTCGCTCACAGAGC 6865
QY 788 GGCTCTATGTCGGGGTCCCATGACTTAACCAAGGCGAGAACTGCGGCTATCGCGGT 847
Db 6866 GGCTTTATATCGGGGTCTCTGACTTAATCAAAAGGCGAGAACTGCGGTTATCGCGGT 6925
QY 848 GCGCGGAGCGGCTGTGACGACTAGCTCGGCTTAATACCTCACATGCTACTTTGAAG 907
Db 6926 GCGCGGAGCGGCTGTGACGACTAGCTCGGCTTAATACCTCACATGCTACTTTGAAG 6985
QY 908 CCGTGTGACGCTGTGAGCTGCGAAGCTTCCAGGACTGACGATGCTCGTGAATGGAGAGC 967
Db 6986 CCGTGTGACGCTGTGAGCTGCGAAGCTTCCAGGACTGACGATGCTCGTGAATGGAGAGC 7045
QY 968 ACCTGTGCTTATCTGTGAAGCGCGGAAACCCAGAGAGCGCGGCGAAGCTTACGAGTCT 1027
Db 7046 ACCTGTGCTTATCTGTGAAGCGCGGAAACCCAGAGAGCGCGGCGAAGCTTACGAGTCT 7105
QY 1028 TCAGGAGGCTATGACTAGGTTACTTGTGCGGCTTGGGACCGCGGCGGCGGCGGCGGCGG 1087
Db 7106 TCAGGAGGCTATGACTAGGTTACTTGTGCGGCTTGGGACCGCGGCGGCGGCGGCGGCGG 7165
QY 1088 ACTTGGAGCTGATAACATCGTGTCTTCTTCAATGTGCTGGTGGCAGACGATGCTTGGCA 1147

Db 7166 ACTTGGAGCTGATAACATCGTGTCTTCTTCAATGTGCTGGTGGCAGACGATGCTTGGCA 7225
QY 1148 AAAGGCTTACTTACTTCAACCGCTGACCCACCGTCCCGCTTGGCGGGCTGCGTGGGAGA 1207
Db 7226 AAAGGCTTACTTACTTCAACCGCTGATCCACCGCTTGGCGGGCTGCGTGGGAGA 7285
QY 1208 CAGCTAGGCACTTCCAGTCAACTCTGCTGCTAGGCAACATCATGATGATGCGGCCACTT 1267
Db 7286 CAGCTAGGCACTTCCAGTCAACTCTGCTGCTAGGCAACATCATGATGATGCGGCCACTT 7345
QY 1268 TGTGGCAAGGATGATCTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1327
Db 7346 TGTGGCAAGGATGATCTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 7405
QY 1328 TTGAAAAGCCCTGGATGTGATGACTTACGCGGCTTGTACTTCCATTTGAGCCACTTTGACC 1387
Db 7406 TTGAAAAGCCCTGGATGTGATGACTTACGCGGCTTGTACTTCCATTTGAGCCACTTTGACC 7465
QY 1388 TACCTCAGATCAITTTGAACGACTCCATGCTTACGCGGCTTGTACTTCCATTTGAGCCACTTTGACC 1447
Db 7466 TACCTCAGATCAITTTGAACGACTCCATGCTTACGCGGCTTGTACTTCCATTTGAGCCACTTTGACC 7525
QY 1448 CAGCTGAGATCAITTTGAACGACTCCATGCTTACGCGGCTTGTACTTCCATTTGAGCCACTTTGACC 1507
Db 7526 CAGCTGAGATCAITTTGAACGACTCCATGCTTACGCGGCTTGTACTTCCATTTGAGCCACTTTGACC 7585
QY 1508 TCTGAGACATCGCGGCGAGAGTGTCTGCGCTAAGTTACTTGTCCAGGGGGGAGGCGCG 1567
Db 7586 TCTGAGACATCGCGGCGAGAGTGTCTGCGCTAAGTTACTTGTCCAGGGGGGAGGCGCG 7645
QY 1568 CCATTTGTGGCAAGTACTCTTCAACCTGGGCGAGTAAAGCAACGCTTTAAACTCTACTCCAA 1627
Db 7646 CCATTTGTGGCAAGTACTCTTCAACCTGGGCGAGTAAAGCAACGCTTTAAACTCTACTCCAA 7705
QY 1628 TTTCCGCTGCTGCTCCGCTGAGACTTGTCCGCGTGGTTCGTTGCTGGCTACAGCGGGGAG 1687
Db 7706 TTTCCGCTGCTGCTCCGCTGAGACTTGTCCGCGTGGTTCGTTGCTGGCTACAGCGGGGAG 7765
QY 1688 ACATATATACAGCTGCTGCTGCGCGGAGCCCG 1721
Db 7766 ACATATATACAGCTGCTGCTGCGCGGAGCCCG 7799

RESULT 12

US-08-384-616-31
; Sequence 31, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 7917 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7862
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..7916
OTHER INFORMATION: /note= "sequence = 1500 - 9416 of
OTHER INFORMATION: SEQ ID NO: 1"
US-08-384-616-31
Query Match 88.5%; Score 1539.6; DB 2; Length 7917;
Best Local Similarity 93.6%; Pred. No. 0; Mismatches 109; Indels 0; Gaps 0;
Matches 1605; Conservative 0;
QY 8 GCATGTCAATGTCCTATACATGACAGCGCCCTGATCACACCGTGCCTGCGGAGGAAA 67
DB 6086 GTGTCTCAATGTCCTACATGACAGCGCCCTGATCACGCCATCGCTGCGGAGGAAA 6145
QY 68 GCAAGCTGCCATCAACCGCGCTGAGCAACTCTTGTGCTGGTGCACCATACCTTGTCTATT 127
DB 6146 GCAAGCTGCCATCAACCGCGTTGAGCAACTCTTGTGCTGCCCAACATAACATGTTTATG 6205
QY 128 CCACAACTCCCGCAGTGAAGCTCGGCGACAGAGAGTCAAGTTCACAGACTGCAAG 187
DB 6206 CCACAACTCTCGCAGCGAGGCTCGGCGACAGAGAGTCAAGTTCACAGACTGCAAG 6265
QY 188 TCTTGACCATCATTTACCGGAGCTGCTCAAGAGATGAAGCGAGGCTCCACAGTGA 247
DB 6266 TCTTGACCAACCACTACCGGAGCTGCTCAAGAGATGAAGCGAGGCTCCACAGTGA 6325
QY 248 AGGCTAAACTGTCTATCTGTAGAAGAGCATGCAAGCTGACGCCCGCGCATTCGGCCAAAT 307
DB 6326 AGGCTAAACTCTTATCCGTAGAGAGAGCTGCAAGCTGAGGCCCGCCACATTTCGGCCAAAT 6385
QY 308 CCAATTTTGGCTATGGGCGAAAGGAGCTCGGAGGCTTATCGAGAGGCGGTTAACCACA 367
DB 6386 CCAAGTTTGGCTATGGGCGAAAGGAGCTCGGAGGCTTATCGAGAGGCGGTTAACCACA 6445
QY 368 TCCGCTCCGTGTGGAGGACTTCTCGGAGGACACTGACACCAATTCAGACCACCATCA 427
DB 6446 TCCACTCCGTGTGGAGGACTTCTCGGAGGACACTGAGACCAATTCAGACCACCATCA 6505
QY 428 TGGCAAAAAATGAGGTTTCTCTGCGTCCAAACAGAGAAAGAGGCGCCGCAAAACCAAGCTCGCC 487

DB 6506 TGGCAAAAAATGAGGTTTCTCTGTGTCAAACAGAGAAAGAGGCGGTATGCCAGCGCCG 6565
QY 488 TCATCGTATTTCCAGACCTGGGAGTTTCGTGTATGCGAGAAAGATGCCCTTTTACGAGTGG 547
DB 6566 TTATCGTATTTCCAGATCTGGGAGTCCGTGTATGCGAGAAAGATGCCCTCTATGATGTGG 6625
QY 548 TTTCCACTCTTCTCAGCCGTGATGGGCTCTCTCATACGGATTCCAAATCTCTCTTAAGC 607
DB 6626 TCTCCACCCCTTCTCAGGTCGTGATGGGCTCTCTCATACGGATTCCAAATCTCTCTTGGGC 6685
QY 608 AGCGGTTCAGTTCCTGTCGAATACCTGCAAGCAAGAAATGCCCTATGGGCTTCTCTCAT 667
DB 6686 AGCGAGTCGAGTTCTCTGGTGAATACCTTGGAAATCAAAAGGAAACCCCATGGGCTTTTTCAT 6745
QY 568 ATGACACCCGCTGTTTGTGACTCAACCGTCACTGAGAAATGACATCCGTGTTCGAGAGTCAA 727
DB 6746 ATGACACTCGCTGTTTCGACTCAACCGTCAACCGTCAACCGTCAACCGTCAACCGTCAAC 6805
QY 728 TTTACCAATGTTGTGACTTTGGCCCCCGAAAGCTAGACAGGCCATAAAGTCTCGTCAAGAGC 787
DB 6806 TTTTACCAATGTTGTGACTTTGGCCCCCGAAAGCTAGACAGGCCATAAATCGTCAAGAGC 6865
QY 788 GGCTCTATGTCGGGGTCCCATGCACTAACTCCAAAGGCGAGAACTCGGGCTATCGCGGT 847
DB 6866 GGCTTTATATCGGGGTCTCTGACTAAATTCAAAAGGCGAGAACTCGGGTATTCGCGGT 6925
QY 848 GCGCGGAGCGCGTCTGACGACTAGCTGGGTAAATACCTTCACATGCTACTCTGAAG 907
DB 6926 GCGCGGAGCGCGTCTGACGACTAGCTGGGTAAATACCTTCACATGCTACTCTGAAG 6985
QY 908 CCGCTGCAGCTGTGAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAATGGAGAGC 967
DB 6986 CCTCTGAGCTGTGAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAACGGAGAGC 7045
QY 968 ACCTTGTGCTTATCTGTGAAAGCGCGGAAACCCAAAGAGAGCGGGAAGCTTCAGAGTCT 1027
DB 7046 ACCTCGTCTTATCTGTGAAAGCGCGGAAACCCAAAGAGAGCGGGAAGCTTCAGAGTCT 7105
QY 1028 TCAGGAGGCTATGACTAGTACTCTGCCCCCTTGGGGACCGGCCCAACCGGAATAAG 1087
DB 7106 TCAGGAGGCTATGACTAGTACTCTGCCCCCTTGGGGACCGGCCCAACCGGAATAAG 7165
QY 1088 ACTTGGAGCTGATAAATCTGTTCTTCCAAATGTCGCTCGACACGATGCAATCTGCGCA 1147
DB 7166 ACTTGGAGCTGATAAATCTGTTCTTCCAAATGTCGCTCGACACGATGCAATCTGCGCA 7225
QY 1148 AAAGGCTGTACTACCTCAACCGTGACCCCAACCGCTTGGCGGGCTGCGTGGGAGA 1207
DB 7226 AAAGGCTGTACTACCTCAACCGTGATCCCAACCGCTTGGCGGGCTGCGTGGGAGA 7285
QY 1208 CAGCTAGGCGACACTCCAGTCAAATCTCTGGGTAGGCAACATCATGATGCGGCCCACTT 1267
DB 7286 CAGCTAGGCGACACTCCAGTCAAATCTCTGGGTAGGCAACATCATGATGCGGCCCACTT 7345
QY 1268 TGTGGGCAAGGATGATTCATGATGACTCACTTCTTCTCAATCTTCTAGCCAGGAGAAC 132
DB 7346 TGTGGGCAAGGATGATTCATGATGACTCACTTCTTCTCAATCTTCTAGCCAGGAGAAC 7405
QY 1328 TTGAAAGCCCTCGATTGTCAGATCTACGGGGCTTGTACTTCCATTCAGCCACTTGAAC 138
DB 7406 TTGAAAGCCCTCGATTGTCAGATCTACGGGGCTTGTACTTCCATTCAGCCACTTGAAC 7465
QY 1388 TACCTTCAGATCAATTCAGGACTCCATGCTTTCAGGCAATTTTACTCCATGATGTTCTC 144
DB 7466 TACCTTCAGATCAATTCAGGACTCCATGCTTTCAGGCAATTTTACTCCATGATGTTCTC 752
QY 1448 CAGGTGAGATCAATAGGGTGGCTTCATGCTTCAGGAAGCTTGGGGTACCACTTGGAG 150
DB 7526 CAGGTGAGATCAATAGGGTGGCTTCATGCTTCAGGAAGCTTGGGGTACCACTTGGAG 758
QY 1508 TCTTGAGACATCGGCGCAGAGAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGCGCG 156

Db 7586 TCTGGAGACATCGGGCCAGAGCGTCCGCTAGGCTACTGTCCAGGAGGAGGCGCG 7645
Qy 1568 CCATTTGTGCGCAAGTACCTCTTCAACTGGCGCAGTAAAGGACCAAGCTTAAACTCACTCCAA 1627
Db 7645 CCACCTGTGCGCAATACCTCTTCAACTGGCGCAGTAAAGGACCAAGCTTAAACTCACTCCAA 7705
Qy 1628 TTCGGCTGCGTCCGCTGGAGCTTGTTCGCGCTGGTTCGCTGCTACAGCGGGGAG 1687
Db 7706 TCCGCTGCGTCCGCTGGAGCTTGTTCGCGCTGGTTCGCTGCTACAGCGGGGAG 7765
Qy 1688 ACATATATACAGCTGCTGTCGTCGCCGACCCCG 1721
Db 7766 ACATATATACAGCTGCTGTCGTCGCCGACCCCG 7799

RESULT 13

US-08-904-686A-31
; Sequence 31, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAWA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLealand, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 31:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 7917 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7862
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..7916
; OTHER INFORMATION: /note= "sequence = 1500 - 9416 of
; OTHER INFORMATION: SEQ ID NO: 1"
US-08-904-686A-31

Query Match 88.5%; Score 1539.6; DB 2; Length 7917;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1605; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 8 GCATGTCAATGCTCTATACATGACGAGCGCCCTGATCACACCGTGGCTCGGAGGAAA 67
Db 6086 GCTGCTCAATGCTCTACACATGGACAGCGCTTGTATCAGCCATGCGCTCGGAGGAAA 6145
Qy 68 GCAAGCTGCCCATCAAGCGCTGAGCAACTCTTCTGCTGGGTACACATAAATCTGCTTAT 127
Db 6146 GCAAGCTGCCCATCAAGCGCTTGTAGCAACTCTTCTGCTGGGTACACATAAATCTGCTTATG 6205
Qy 128 CCACAACATCCGCGAGTGCAGCCTGGGAGAGAGGTGACCTTTGACAGACTGCAAG 187
Db 6206 CCACAACATCTCGCAGCGCAGGCTGCGGAGAGAGGTGACCTTTGACAGACTGCAAG 6265
Qy 188 TCCTGGAGCATCATTTACCGGAGCGTGTCAAGGAGATGAAGGCGAAGCGTCCACAGTGA 247
Db 6266 TCCTGGAGCATCATTTACCGGAGCGTGTCAAGGAGATGAAGGCGAAGCGTCCACAGTGA 6325
Qy 248 AGGCTAAACTCTATCTGTAGAGAGCATGCAAGCTGACGCGCCCGCCCATTTGGGCCAAAT 307
Db 6326 AGGCTAAACTCTATCTGTAGAGAGCATGCAAGCTGACGCGCCCGCCCATTTGGGCCAAAT 6385
Qy 308 CCAAAATTTGGCTATGGGGGAAAAGAGCGTCCGAGCGCTTATCCAGAGGCGGTTAACACA 367
Db 6386 CCAAAATTTGGCTATGGGGGAAAAGAGCGTCCGAGCGCTTATCCAGAGGCGGTTAACACA 6445
Qy 368 TCAGCTCCGTGTGAGAGGACTTCTGAGAGGAGCTGACACACCAATTCAGACCAACATCA 427
Db 6446 TCAGCTCCGTGTGAGAGGACTTCTGAGAGGAGCTGACACACCAATTCAGACCAACATCA 6505
Qy 428 TGGCAAAAATGAGGTTTCTGCTGCCAAGAGAGGAGCGGCGCAAAACAGCTCGCC 487
Db 6506 TGGCAAAAATGAGGTTTCTGCTGCCAAGAGAGGAGCGGCGGCGCAAAACAGCTCGCC 6565
Qy 488 TCATCGTATTTCCAGACCTGGGAGTTGTTGATGCGAGAGAGTGGCCCTTTTACGACGTGG 547
Db 6566 TTATCGTATTTCCAGACCTGGGAGTTGTTGATGCGAGAGAGTGGCCCTTTTACGACGTGG 6625
Qy 548 TTTCCACTCTTCTCAGCGCGTGTATGGGCTCTCTATACGAGTTCCAAATATCTCTCTAAGC 607
Db 6626 TTTCCACTCTTCTCAGCGCGTGTATGGGCTCTCTATACGAGTTCCAAATATCTCTCTAAGC 6685
Qy 608 AGCGGGTCGAGTTCCCTGTAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 667
Db 6686 AGCGGGTCGAGTTCCCTGTAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 6745
Qy 668 ATGACACCCGCTGTTTGTACTCAACGGTCACTGAGAAATGACATCCGTTTGGAGGTCAA 727
Db 6746 ATGACACTCGCTGTTTGTACTCAACGGTCACTGAGAAATGACATCCGTTTGGAGGTCAA 6805
Qy 728 TTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCGCATTAAGTCTGCTCACAGC 787
Db 6806 TTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCGCATTAAGTCTGCTCACAGC 6865
Qy 788 GGCTCTATGTCGGGGGTCCCATGACTAACTCCAAAGGGGAGAACTGCGGGCTATCGCGGT 847

Db 6866 GCGTTTATATCGGGGTCCTCTGACTAAATCAAAGGCGAGAACTGCGGTATTCGCGGT 6925
Qy 848 GCGCGGAGCGGGTGTGAGACTAGCTAGCTGCGGTATATACCTCACATGCTACTTGAAGG 907
Db 6926 GCGCGGAGCGGGTGTGAGACTAGCTAGCTGCGGTATATACCTCACATGCTACTTGAAGG 6985
Qy 908 CCGCTGCAGCTGTGAGCTGCCAAGCTCCAGGACTCCAGGACTCCAGGACTCCGTAATGGAGACG 967
Db 6986 CCTCTGCAGCTGTGAGCTGCCAAGCTCCAGGACTCCAGGACTCCAGGACTCCGTAATGGAGACG 7045
Qy 968 ACCTTGCTGTTATCTGTGAAGCGCGGAACCCAAAGAGAGCGCGGCAAGCTTACGAGTCT 1027
Db 7046 ACCTGCTGTTATCTGTGAAGCGCGGAACCCAAAGAGAGCGCGGCAAGCTTACGAGTCT 7105
Qy 1028 TCAGGAGGCTATGACTAGGTACTTCTGCCCTCCCTGGGGACCCGCCCAACCGGAATACG 1087
Db 7106 TCAGGAGGCTATGACTAGGTACTTCTGCCCTCCCTGGGGACCCGCCCAACCGGAATACG 7165
Qy 1088 ACTTGGAGCTGATTAACATCGTGTCTTCCATATGTGTGGTGGACACGATGCTCTGGCA 1147
Db 7166 ACTTGGAGCTGATTAACATCATGTTCTTCCATATGTGTGGTGGACACGATGCTCTGGCA 7225
Qy 1148 AAAGGCTGACTACTCACCCTGACCCACCCGTCCTTGGCGGGCTGCGTGGGAGA 1207
Db 7226 AAAGGCTGACTACTCACCCTGATCCACCCCTAGCAGGGCTGCGTGGGAGA 7285
Qy 1208 CAGCTTAGCACACTCCAGTCAACTCTTGGCTAGGCAACATCATCATGTATGCGGCCACTT 1267
Db 7286 CAGCTTAGCACACTCCAGTCAACTCTTGGCTAGGCAACATCATCATGTATGCGGCCACTT 7345
Qy 1268 TGTGGGCAAGGATGATCTGATGACTCACTTCTTCTCCATCTTCTAGCCGAGGAGAAC 1327
Db 7346 TGTGGGCAAGGATGATCTGATGACTCACTTCTTCTCCATCTTCTAGCCGAGGAGAAC 7405
Qy 1328 TTGAAAAAGCCCTGATGTGATGATCTAGGGGCTGTGTACTTCAATGAGCACTTGACC 1387
Db 7406 TTGAAAAAGCCCTGACTGCCAGATCTACGGGGCTGTGTACTTCAATGAGCACTTGACC 7465
Qy 1388 TACCTCAGATCAATCAAGACTCCATGTCTTAGGGCACTTTTCACTCCATAGTTACTCTC 1447
Db 7466 TACCTCAGATCAATCAAGACTCCATGTCTTAGGGCACTTTTCACTCCATAGTTACTCTC 7525
Qy 1448 CAGGTGAGATCAATAGGTTGGCTTCATGCTCAGGAAGCTTGGGTACCACTTGGGAG 1507
Db 7526 CAGGTGAGATCAATAGGTTGGCTTCATGCTCAGGAAGCTTGGGTACCACTTGGGAG 7585
Qy 1508 TCTGGAGACATCGGCGCAGAAAGTGTCCGGCTTAAGTTACTGTCCAGGGGGGAGGGCG 1567
Db 7586 TCTGGAGACATCGGCGCAGAGAGCTCCGGCTTAGGCTACTGTCCAGGGGGGAGGGCG 7645
Qy 1568 CATTGTGCGAAGTACTCTTCACTGGGAGTAAAGCAAGCTTAACTCACTCCAA 1627
Db 7646 CCACTTGTGGCAATACCTCTTCACTGGGAGTAAAGCAAGCTTAACTCACTCCAA 7705
Qy 1628 TTCCGCTCGCTCCGGCTGGACTTGTCCGGCTGTCTGCTGCTGCTACAGCGGGGAG 1687
Db 7706 TCCGGCTCGCTCCGGCTGGACTTGTCCGGCTGTCTGCTGCTGCTTACAGCGGGGAG 7765
Qy 1688 ACATATATCACAGCTGTCTCGTGGCCGACCCCG 1721
Db 7766 ACATATATCACAGCTGTCTCGTGGCCGACCCCG 7799

RESULT 14

US-09-315-850-31
; Sequence 31, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa

APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
ADDRESSEE: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,850
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686
FILING DATE: 01-AUG-1997
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mcleland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 7917 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7862
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..7916
OTHER INFORMATION: /note= "sequence = 1500 - 9416 of
OTHER INFORMATION: SEQ ID NO: 1"
US-09-315-850-31

Query Match 88.5%; Score 1539.6; DB 3; Length 7917;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1605; Conservative 0; Mismatches 109; Indels 0; Gaps 0

QY 8 GCATGTCAATCTCTTATACATGGACGGCCCTGATCACACCTGCGCTGGCGAGGAA 67

Db 6086 GCTGCTCAATGTCTACATGAGCAGAGCGCGCTTGATCAGCCATGCGCTCGGAGGAAA 6145
Qy 68 GCAAGCTGCCATCAACGGCTGAGCAACTCTTGTGCTGGTCAACATACCTGGTCTATT 127
Db 6146 GCAAGCTGCCATCAACGGCTGAGCAACTCTTGTGCTGGCACCACCAATCAATCATGTTATG 6205
Qy 128 CCACAAATCCCGCAGTCAAGCTTCGCGCAGAGAGGTCACTTTTGACAGACTGCAAG 187
Db 6206 CCACAAATCTGCGAGCAGCGCTCGCGCAGAGAGGTCACTTTTGACAGACTGCAAG 6265
Qy 188 TCTGAGCAGATCAATACCGGAGCGTCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA 247
Db 6266 TCTGAGCAGCAGTCACTACCGGAGCGTCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA 6325
Qy 248 AGGCTAACTGCTATCTGTAGAGAGCATGCAAGTGAAGCGCGCGGAGTTCGGCAAAAT 307
Db 6326 AGGCTAACTGCTATCTGTAGAGAGCATGCAAGTGAAGCGCGCGGAGTTCGGCAAAAT 6385
Qy 308 CCAAAATTTGGCTATGGGCAAGGAGCTCCGAGGCTTATCCAGAGGCGGCTTACACCA 367
Db 6386 CCAAGTTTGGCTATGGGCAAGGAGCTCCGAGGCTTATCCAGAGGCGGCTTACACCA 6445
Qy 368 TCGCTCCGCTGGGAGGACTTGTGGAGGACACTGACACCAATTCAGACCAACATCA 427
Db 6446 TCACTCCGCTGGGAGGACTTGTGGAGGACACTGAGACCAATTCAGACCAACATCA 6505
Qy 428 TGGCAAAAATGAGGTTTTCTCGCTCCAAACAGAGAGAGGCGCGCAAAACAGCTCGCC 487
Db 6506 TGGCAAAAATGAGGTTTTCTGTGTCCAAACAGAGAGAGGCGCGTAAGCAGCGCGCC 6565
Qy 488 TCATCGTATTCACAGACTGGGAGTTGCTGTATGCGAGAGATGCGCCCTTACAGAGTGG 547
Db 6566 TTATCGTATTCACAGACTGGGAGTTGCTGTATGCGAGAGATGCGCCCTTATGATGTGG 6625
Qy 548 TTTCCACTCTTCTCAGCGCGTATGAGGCTCTCATACGATTCCAATACTCTCCTAAGC 607
Db 6626 TCTCCACCCCTCTCAGGTCGTGATGGGCTCTCATACGATTCAGGATCTCTCTCGGCG 6685
Qy 608 AGCGGTCGAGTTCTCGTGTGAATACCTGGAAAGCAAGAAATGCCCTATGGCTTCTCAT 667
Db 6686 AGCGAGTCGAGTTCTCGTGTGAATACCTGGAAATCAAGAAANACCCCATGGGCTTTTCAT 6745
Qy 668 ATGACACCGCGTGTGTTGACTCAACCGTCACTGAGATGACATCCGCTGTGAGGAGTCAA 727
Db 6746 ATGACACTCGCTGTGTTGACTCAACCGTCAACCGTCAACGAGAACGACATCCGTTGAGGAGTCAA 6805
Qy 728 TTTACCAATGTTGACTTGGCCCCCGAGCTAGACAGGCCATAAGTTCGCTCAAGAGC 787
Db 6806 TTTACCAATGTTGACTTGGCCCCCGAGCGCAGACAGGCCATAAAATCGCTCAAGAGC 6865
Qy 788 GGCTCTATGTCGGGGTCCCATGACTTAATCCAAAGGCGAGACTGCGGCTATCGCGGT 847
Db 6866 GGCTTTATATCGGGGTCTCTGACTTAATTCAAAAGGCGAGAACTGCGGTTATCGCGGT 6925
Qy 848 GCGCGCGAGCGGCTGTGACGACTAGCTGCGGTAATACCTCACATGCTACTTGAAGG 907
Db 6926 GCGCGCGAGCGGCTGTGACGACTAGCTGCGGTACACCTCACATGTTACTTGAAGG 6985
Qy 908 CGGCTGCAGCCTGTGAGCTGCCAAGCTCCAGGACTGCAACGATGCTCGTGAATGAGAGC 967
Db 6986 CCTCTGCAGCCTGTGAGCTGCGAAGCTCCAGGACTGCGACGATGCTCGTGAAGGAGAGC 7045
Qy 968 ACCTGTGCTTATCTGTGAAGCGCGGAAACCCAGAGAGAGCGCGGAGCGCTACGAGTCT 1027
Db 7046 ACCTGTGCTTATCTGTGAAGCGCGGAAACCCAGAGAGAGCGCGGAGCGCTACGAGTCT 7105
Qy 1028 TCACGAGGCTATGACTAGTACTCTGCCCCCTCGGAGCGCGGCGGAGCGCGGAGCGCTACGAGTCT 1087
Db 7106 TCACGAGGCTATGACTAGTACTCTGCCCCCTCGGAGCGCGGAGCGCGGAGCGCTACGAGTCT 7165
Qy 1088 ACTTGGAGCTGATAACATCGTGTCTCTCAATGTGCTGCGACACGATGCTATCGGCA 1147
Db 7166 ACTTGGAGCTGATAACATCAATGTTCTCTCATGTGCTGCTGCCACCATGCTACGCA 7225

Qy 1148 AAAGGTTGTTACTACTCTACCCCGTGAACCCACCGCTCCCTTTCGCGGCGCTGCTGGAGAGA 1207
Db 7226 AAAGGTTGTTACTACTCTACCCCGTGAACCCACCGCTCCCTTTCGCGGCGCTGCTGGAGAGA 7285
Qy 1208 CAGCTAGGACACATCCAGTCAACTCTCGCTAGGCAACATCATATGATATGCGGCCACTT 1267
Db 7286 CAGCTAGGACACATCCAGTCAACTCTCGCTAGGCAACATCATATGATATGCGGCCACTT 7345
Qy 1268 TGTGGCAAGGATGATTTCTGATGACTCACTTTCTTCTCCATCTTCTAGCCAGGAGCAAC 1327
Db 7346 TGTGGCAAGGATGATTTCTGATGACTCACTTTCTTCTCCATCTTCTAGCCAGGAGCAAC 7405
Qy 1328 TTGAAAAAGCCCTGGATGTCAGATCTACGCGGCTTGTACTCCATTGAGCACTTGACC 1387
Db 7406 TTGAAAAAGCCCTGGATGTCAGATCTACGCGGCTTGTACTCCATTGAGCACTTGACC 7465
Qy 1388 TACCTCAGATCATTAAGACGACTCCATGCGCTTACGCGATTTTCACTCCATGATTTACTCTC 1447
Db 7466 TACCTCAGATCATTAAGACGACTCCATGCGCTTACGCGATTTTCACTCCATGATTTACTCTC 7525
Qy 1448 CAGGTGAGATCAATAGGTTGGCTTCATGCGCTCAGGAAGCTTGGGTACACCTTTCGAG 1507
Db 7526 CAGGTGAGATCAATAGGTTGGCTTCATGCGCTCAGGAAGCTTGGGTACACCTTTCGAG 7585
Qy 1508 TCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAAGTTACTGTCCAGGGGGAGGGCG 1567
Db 7586 TCTGGAGACATCGGGCCAGAGCGTCCGCGCTAGGCTACTGTCCAGGGAGGGAGGGCG 7645
Qy 1568 CCATTGTTGGCAAGTACTCTTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCTACTCAA 1627
Db 7646 CCACTTGTGGCAAAATACCTCTTCAACTGGGCGAGTAAAAAACCAAACTTAAACTCTACTCAA 7705
Qy 1628 TTTCCGGCTGCGTCCGCGTGGACTTGTCCGCGTGGACTTGTCCGCGTGGTTCGTTGCTGCTACAGCGGGAG 1687
Db 7706 TTTCCGGCTGCGTCCGCGTGGACTTGTCCGCGTGGACTTGTCCGCGTGGTTCGTTGCTGCTACAGCGGGAG 7765
Qy 1688 ACATATATACAGCGCTGTCTGTCGCCGACCCCG 1721
Db 7766 ACATATATACAGCGCTGTCTGTCGCCGACCCCG 7799

RESULT 15

US-08-324-977-13
; Sequence 13, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
; ADDRESS: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466

; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9030 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..9030
; OTHER INFORMATION: /note: "sequence = 333 - 9362 of
; OTHER INFORMATION: SEQ ID NO: 1"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..9030
; US-08-324-977-13

Query Match 88.5%; Score 1539.6; DB 1; Length 9030;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1605; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy	8	GCATGTCAATGTCTTATACATGGACAGCGCCCTTGATCACACCGTGGCTGCGGAGGAAA	67
Db	7253	GCTGCTCAATGCTCTACACATGGACAGCGCGCTTGATCACGCCATGCGTGGCGAGGAAA	7312
Qy	68	GCAAGCTGCCATCAACCGCTGAGCAACTCTTGCTGCGTCAACATAACCTGTCTATT	127
Db	7313	GCAAGCTGCCATCAACCGCTGAGCAACTCTTGCTGCGGCAACATAACATGCTTTATG	7372
Qy	128	CCACAACATCCGGAGTCAAGCCTGCGGAGAGAGAGGTCAACCTTTGACAGACTGCAAG	187
Db	7373	CCACAACATCTCGAGCGCAGCGCTGCGGAGAGAGAGGTCAACCTTTGACAGACTGCAAG	7432
Qy	188	TCCTGGAGCATCAATTACCGGAGCATGTCTCAAGGAGATGAAGCGGAAGCGTCCACAGTGA	247
Db	7433	TCCTGGAGCATCAATTACCGGAGCATGTCTCAAGGAGATGAAGCGGAGCGTCCACAGTTA	7492
Qy	248	AGGCTAAATCTGCTATCTGTAGAAGAGCATCAAGCTGACGCCGCCCGCATTTGCGCCAAAT	307
Db	7493	AGGCTAAATCTGCTATCTGTAGAAGAGCATCAAGCTGACGCCGCCCGCATTTGCGCCAAAT	7552
Qy	308	CCAAATTTGGCTATGGGCAAGAGCATGTCGGAGCCTATCCAGCAGGCGCGTTAAACACA	367
Db	7553	CCAAATTTGGCTATGGGCAAGAGCATGTCGGAGCCTATCCAGCAGGCGCGTTAAACACA	7612
Qy	368	TCCGCTCCGTGTGGAAGGACTTGCTGGAGGACACTGACACACCAATTCAGACCAACCATCA	427
Db	7613	TCCACTCCGTGTGGAAGGACTTGCTGGAGGACACTGTCGACACCAATTCAGACCAACCATCA	7672

Qy	428	TGCGAAAAATGAGGTTTTTCTGGTCCAAACGAGAAAGAGGCGCCAAACCAAGCTCGCC	487
Db	7673	TGCGAAAAATGAGGTTTTTCTGTGTCAAACGAGAGAAAGAGGCGCGTAAGCAGGCGCGC	7732
Qy	488	TCATCGTATTCCAGACCTGGGAGTTCTGTATGCGAGAAGATGGCCCTTTACGACGTGG	547
Db	7733	TTATCGTATTCCAGATCTGGGAGTCCGTGTATGCGAGAAGATGGCCCTCTATGATGTGG	7792
Qy	548	TTTCCACTCTTCTCAGGCGTGATGGGCTCTCTATACGGAATTCGAATATCTCTCTAAGC	607
Db	7793	TCCTCACCTCTCTCAGGTCGTATGGGCTCTCTATACGGAATTCGAATATCTCTCTGGGC	7852
Qy	608	AGCGGTGAGTTCTCTGGTGAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT	667
Db	7853	AGCGAGTCGAGTTCTCTGGTGAATACCTGGAAATCAAGAAAGCAAGAAATGCCCTATGGGCTTCTCAT	7912
Qy	668	ATGACACCCGCTGTTTTGACTCAACGCTCACTGAGAAATGACATCCGTGTTGAGGACTCAA	727
Db	7913	ATGACACTCGCTGTTTTGACTCAACGCTCAACGCTCAACGAGAACACATCCGTGTTGAGGACTCAA	7972
Qy	728	TTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCCATAAGTTCGCTCACAGAGC	787
Db	7973	TTTACCAATGTTGTGACTTTGGCCCCCGAAGCCAGACAGGCCATAAATCGCTCACAGAGC	8032
Qy	788	GGCTCTATGTCGCGGGTCCCATGACTAACTCCAAAGGCGAGAACTGCGGCTATCGCCGT	847
Db	8033	GGCTTTATATCGGGGTCCTCTGACTAATTCAAAAGGCGAGAACTGCGGCTTATCGCCGT	8092
Qy	848	GCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAAATACCTCACATGCTACTTGAAGG	907
Db	8093	GCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAAACACCTCTACATGTTACTTGAAGG	8152
Qy	908	CCGCTCGAGCTGTCAGCTGCCAGCTCCAGACTGCACAGCTGCTGTAATGGAAGC	967
Db	8153	CCTCTGAGCCTGTCAGCTGCGAGCTCCAGACTGCGACGATGCTGTAAGCGAGAGC	8212
Qy	968	ACCTTGTCTGTTATCTGTGAAAGCGCGGAAACCCCAAGAGGACGCGGAGCTTACGAGTCT	1027
Db	8213	ACCTCGTCTTATCTGTGAAAGCGCGGAAACCCCAAGAGGACGCGGAGCTTACGAGTCT	8272
Qy	1028	TCAGGAGCTATGACTAGGTACTCTGCCCCCTTGGGAGCCGCCCCCAACCGGAATACG	1087
Db	8273	TCAGGAGCTATGACTAGGTACTCTGCCCCCTTGGGAGCCGCCCCCAACCGGAATACG	8332
Qy	1088	ACTTGAGCTGATAACATCGTGTTCCTCCANTGTGTCGGTTCGACACGATGATCTCGCA	1147
Db	8333	ACTTGAGCTGATAACATCATGTTCCTCCAATGTGTCGGTTCGCCCCACGATGCAATCAGGCA	8392
Qy	1148	AAAGGTGTACTACCTCACCGGTGACCCCAACCGTCCCCCTTGGCGGGGTGCGTGGGAGA	1207
Db	8393	AAAGGTGTACTACCTCACCGGTGATCCCAACCCGCCCTTAGCACCGGCTGCGTGGGAGA	8452
Qy	1208	CAGCTAGGCACACTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGCGCCCACTT	1267
Db	8453	CAGCTAGACACTCCAGTTACTCTCTGGCTAGGCAACATTAATATGATGCGCCCACTT	8512
Qy	1268	TGTGGCAAGGATGATCTGATCACTTCTTCTTCCATTCCTTCTTAGCCAGGAGCAAC	1327
Db	8513	TGTGGCAAGGATGATCTGATCACTTCTTCTTCCATTCCTTCTTAGGAGGAGCAAC	8572
Qy	1328	TTGAAAAGCCCTGGATGTGATGATCTAGGGGCTTTACTCCATTTAGGCCACTTTGACC	1387
Db	8573	TTGAAAAGCCCTGGATGTGATGATCTAGGGGCTTTACTCCATTTAGGCCACTTTGACC	8632
Qy	1388	TACTCAGATCAATTGAAAGCACTCCATGGTCTTAGCGCAATTTTCACTCCATAGTTACTCTC	1447
Db	8633	TACTCAGATCAATTGAAAGCACTCCATGGCTTAGCGCAATTTTCACTCCATAGTTACTCTC	8692
Qy	1448	CAGGTGAGATCAATAGGCTGGCTTTCATGTGCTCAGGAAGCTTGGGGTACCAACCTTCGGAG	1507
Db	8693	CAGGTGAGATCAATAGGCTGGCTTTCATGTGCTCAGGAAGCTTGGGGTACCAACCTTCGGAG	8752
Qy	1508	TCTGGAGACATCGGCGCAGAAAGTGTGCGGCTTAAGTTACTGTCCAGGGGGGAGGGCGC	1567

D _b	8753	TCTGGAGACATCGGCGCCAGAGCGTCCGCGCTAGGCTACTGTCCCAGGAGGGGCGG	8812
Q _y	1568	CCATTTCGGCAAGTACTCTTTAACTGGGCGTAGGAACAAGCTTTAAACTCACTCCAA	1627
D _b	8813	CCACTTCGTGGCAATAAGCTCTTCAACTGGGCAGTAAAAACCAAACCTTTAAACTCACTCCAA	8872
Q _y	1628	TTCCGGCTGCGTCCCGGCTGGACTTTGTCCCGCTGGTTTCGTTGCTTGCTACAGCGGGGAG	1687
D _b	8873	TCCCGGCTGGCTCCCGCTGAGCTTTGTCCCGCTGGTTTCGTTGCTTGCTACAGCGGGGAG	8932
Q _y	1688	ACATATATCACAGCTGCTCTCGTGCCCGACCCCG	1721
D _b	8933	ACATATATCACAGCTGCTCTCGTGCCCGACCCCG	8966

Search completed: October 3, 2005, 22:27:18
Job time : 303.103 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 09:28:05 ; Search time 7968.72 Seconds
(without alignments)
10963.461 Million cell updates/sec

Title: US-10-712-479-1

Perfect score: 1803

Sequence: 1 atggctagcatgcaatgc.....ggcatcaccatcaccatcac 1803

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1642.6	91.1	9547	14	AB049091 Hepatitis
2	1640.4	91.0	9374	14	AF207753 Hepatitis
3	1639.4	90.9	9548	14	D89815 Hepatitis C
4	1628.8	90.4	9379	14	AF165059 Hepatitis
5	1628.2	90.3	9379	14	AF165060 Hepatitis
6	1628	90.3	1773	14	AB057599 Hepatitis
7	1619.2	89.8	9616	14	AB049088 Hepatitis
8	1618.6	89.8	9377	14	AF207763 Hepatitis
9	1617.8	89.7	9375	14	AF207767 Hepatitis
10	1615.4	89.6	9379	14	AF207754 Hepatitis
11	1613.8	89.5	9456	14	HPCRNA
12	1607.4	89.2	9379	14	AF165062 Hepatitis C
13	1607.4	89.2	9414	14	HPCUNKCD8
14	1607.4	89.2	9415	14	HCU16362
15	1607.4	89.2	9448	14	UI6362 Hepatitis C
16	1607.4	89.2	9460	14	DI3558 Hepatitis C
17	1607.4	89.2	9472	6	AR301390
18	1605.6	89.1	1773	14	AB057601
19	1605.2	89.0	9535	14	D85516 Hepatitis C

ALIGNMENTS

RESULT 1
AB049091
LOCUS AB049091 9547 bp RNA linear VRL 22-AUG-2002
DEFINITION Hepatitis C virus gene for polyprotein, complete cds,
isolate:HCV142.
ACCESSION AB049091.1 GI:11559448
VERSION
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K.,
Hatahara, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and
Mishiro, S.
TITLE Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
with hepatocellular carcinoma: the 'progression score' revisited
JOURNAL Hepatol. Res. 20 (2), 161-171 (2001)
PUBMED 11348851
REFERENCE 2 (bases 1 to 9547)
AUTHORS Mishiro, S.
TITLE Direct Submission
JOURNAL Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,
Department of Medical Sciences, 6-3-22 Higashi Oh-1, Shinagawa-Ku,
Tokyo 140-8522, Japan (E-mail: shunji.mishiro@po.toshiba.co.jp,
Tel: 81-3-3764-8981, Fax: 81-3-3764-8992)
FEATURES
source location/Qualifiers
1..9547
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="HCV142"
/specific_host="human"
/db_xref="taxon:11103"
/tissue_type="serum"
/note="HCV-genotype: 1b"
<1..259
/note="partial"
260..9292
/codon_start=1
/product="polyprotein"
/protein_id="BAB18804.1"

AB080299 Hepatitis
AB031322 Hepatitis
AB049093 Hepatitis
AF207756 Hepatitis
AB057680 Hepatitis
AF054250 Hepatitis
AF207772 Hepatitis
AB057602 Hepatitis
AB049090 Hepatitis
AF165058 Hepatitis
AF207758 Hepatitis
U01214 Hepatitis C
AB057688 Hepatitis
AF313916 Hepatitis
AF333324 Hepatitis
AF165061 Hepatitis
AR119832 Sequence
CQ819761 Sequence
AF054247 Hepatitis
AX593183 Sequence
AF207770 Hepatitis
D10750 Hepatitis C
AF054248 Hepatitis
AX593179 Sequence
AX593184 Sequence
AY044867 Hepatitis

7572	CCAAGCTGCCCATCAACCGCGCTGAGCAACTCTCTGCTGCGTACACACAAATCGTGGTATG	128	CCACCAATCCCGCAGTGAAGCTCTCGGCAGAAAGAGTCACTTTTGACAGACTGCAAG	197	GCATCTG
------	---	-----	---	-----	--

7572	GCAAGCTGCCCATCAACCGCTGAGCAACTCTCTGCTGCGTCAACCAACATGGTGTATG	7631
128	CCACAAATCCCGCAGTGAAGCCTCGGCAGACAGAGAGGTCACTTTTGACAGACTGCAAG	187
7632	CCACAAATCCCGCAGCAGCAAGCCTCGGCAGACAGAGAGGTCACTTTTGACAGACTGCAAG	7691
188	TCCTGCACCATCATTAACCGGAGCTGCTCAAGGAGATGAAGCGCAAGGGGTCCACAGTGA	247
7692	TCCTGCACCATCACTACCGGAGCGTCTCAAGGAGATGAAGCGCAAGGGGTCCACAGTGA	7751
248	AGGCTAAATCTGCTATCTGTAGAAGAAAGCATGCAAGCTGACGCCCCCGCATTTGGCCAAAT	307
7752	AGGCTAAATCTCTATCTGTAGAAGAAAGCATGCAAGCTGACGCCCCCGCATTTGGCCAAAT	7811
308	CCAAATTTGGCTATGGGGGMAAGAGAGCTCGGAGGCTTCCAGCAGGGGCGGTAAACCACA	367
7812	CCAAATTTGGCTATGGGGGMAAGAGAGCTCGGAGGCTTCCAGCAGGGGCGGTAAACCACA	7871
368	TCCGCTCCGTGTGGNAGGACTTCTCTGGAGGACACTGCACACACCAATTCAGACCACCATCA	427
7872	TCCGCTCCGTGTGGNAGGACTTCTCTGGAGGACACTGCACACACCAATTCAGACCACCATCA	7931
428	TGGCAAAAAATGAGGTTTTCTGCGTCCAAACGACAGAAAGGAGGCCGCAACACAGCTGCGC	487
7932	TGGCAAAAAATGAGGTTTTCTGCGTCCAAACGACAGAAAGGAGGCCGCAACACAGCTGCGC	7991
488	TCATCGTATTTCCACAGACTGGGAGTTCGTGTATGCGAGAAAGATGGCCCTTTACAGCGTGG	547
7992	TTATCGTATTTCCACAGACTGGGAGTTCGTGTATGCGAGAAAGATGGCCCTTTATGACGTAG	8051
548	TTTCCACTCTTCTCCTCAGGCGTGTATGGGCTCTCTATACGATTCGCAATCTCTCTAAAGC	607
8052	TTTCCACTCTTCTCCTCAGGCGTGTATGGGCTCTCTATACGATTCGCAATCTCTCTAAAGC	8111
608	AGCGGTCGAGTCTCTCGTGAATACCTTGGAAAGCAAAGAAATGCCCTATGGGCTTCTCAT	667
8112	AGCGGTCGAGTCTCTCGTGAAGACCTTGGAAATCAAAGAAAGTGGCTTATGGGCTTCTCAT	8171
668	ATGACACCCCGCTGTTTTGACTCAACCGGTCACTGAGAAATGATCATCCGTTTGAAGAGTCAA	727
8172	ATGACACCCCGTGTGTTTTGACTCAACCGGTCACTGAGAAATGATCATCCGTTTGAAGAGTCAA	8231
728	TTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCCATTAAGTTCGCTCACAGAGC	787
8232	TTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCCATTAAGTTCGCTCACAGAGC	8291
788	GGCTCTATCTCGGGGTCCTATGACTCAACCGGTCACTGAGAAATGATCATCCGTTTGAAGAGTCAA	847
8292	GGCTCTATCTCGGGGTCCTATGACTCAACCGGTCACTGAGAAATGATCATCCGTTTGAAGAGTCAA	8351
848	GCCGCGCAGCGGGGTGCTGACGACTAGCTCGGCTTAATACCTCTCACATGCTACTTTGAAG	907
8352	GCCGCGCAGCGGGGTGCTGACGACTAGCTCGGCTTAATACCTCTCACATGCTACTTTGAAG	841
908	CCGCTGACAGCCTGTCGAGCTTCCAAAGCTTCCAGGATGTCACGATGCTGCGTGAACGAGAGC	967
8412	CCACTGCGAGCCTGTCGAGCTGGAAGCTCCAGGATGTCACGATGCTGCGTGAACGAGAGC	847
968	ACCTTGTGCTTATCTGTGAAGCGCGGGAACCCACAGAGAGCGCGCAAGCCTACGAGTCT	102
8472	ACCTTGTGCTTATCTGTGAGAGCGCGGGAACCCACAGAGAGTGGCGCAAGCCTACGAGTCT	853
1028	TCACGAGCCTATGACTAGGTACTCTGCCCCCTCGGAGCCCGCCCAACCGGAATACG	108
8532	TCACGAGCCTATGACTAGGTACTCTGCCCCCTCGGAGCCCGCCCAACCGGAATACG	859
1088	ACTTGGAGCTGATAAATCGTGTCTCTCCAATGTGTCTGCTGCGACACGATGCAATCTGGCA	114
8592	ACTTGGAGCTGATAAATCGTGTCTCTCCAATGTGTCTGCTGCGACACGATGCAATCTGGCA	865
1148	AAAGGCTGTACTACCTCAACCCGCTGACCCGCTGACCCGCTTGGCGGGGCTGGCTGGAGA	120
8652	AAAGGCTGTACTACCTCAACCCGCTGACCCGCTGACCCGCTTGGCGGGGCTGGCTGGAGA	871

/product="polypeptide"		/protein_id="BAA25076.1"	
/translation="MSTNPKPQKTKRNTNRRPDVKKPGGQIVGVVLLPRRPRRLGVRATKTSERSQPRRQPIKARRPGRSTWAOQGYWPPLVYNGEGLGAGMLSPRGSPKMGNDPRRSNGLKVIDTLTCGFADLMGYIPLVGPALGGARALAHGVRLDGVNATGNLPKGSFSLFLLALUSLCTIPASAYEVNRVSGIYHVTNDCSNSSIVBAADVIMHAPGCVPCVENNRSSCWALPTLAARNASVPTTLRRHVDLLVGTAAFSAMYSOLLGCSVLIQLPTFPFRHETVQDCNCSYPGHVSHRMWMMNWSPTAALVVSQLRIPQAVMDVAGAHGVLAGLAIYSWGNWAKVLIWMLFAGVVGHRVTVGGVQGHVSTLSPRGAQSKQIOLVNTNGSHRINTALNCDSLKITGFLAALFYTHKFNASGCPERMASCRSDKEDQNGPITYAQNDSORPCYHAPROCCGIVPASQVCGPVYCFPSPVVGTTDRFCAPYINMGDNTDLLNTRPHGNWFGCTMNSSTGFTKTCGGPFCNTRGNNTLTCTPCRCPHDAITTKCGSPMLTPRCLVDVYHMLPVPCVNFPIFKVMYVGVGHEHRLDAACNWRGERCDLDRDAELSPULLSTFEWQLIPCSYITLIPALSTLHLHONIVDIOVLXIGISAVUSIAIKWEYVVLPLLADARVCAKLMULLIAOEAALLENVLVNAASVVGNGHMLPFMFCCAAWYMKGRVLPQAAFYGVWPLLLALLAPRAYADREWASCGGVFVGLALLTLSPYKVFLEARIIMWLYFITKAEHALLOVSLPVRNGRDAIILLMCVHPELIIDITKILLSLILGLMVLQGLRDLRVFVFRAGQLIRACMLKKAAGHYVQMAFVKLAALTGTVYVHUTPLQDWAHVGLRDLAVNVPVFSAMETKVTWAGDAAACGDIISGLPVSAARGKEILLGPAQSFQCGWRLLAPITAVSQOTRGLGCIITSLGRDKNQVEGVVSTAKQSFPLATCVNGACVTFHAGSKTLARAKGPIITOMYINVDODLVGPAPPGARSLTPTCGSSDLXLVLRHADYIPVRRGDSRGLSPRPIISLYKSGSGGLPCPSGHVVGIIFFRAAVCTRGVAKAVDIPVESMETLMRSVFTDNSTPPAVPQFVAHLHAPTGGSKTKVPAAYAAQQVWVILNPSVAATLFGFAYMSKAGHDNIRTVGRTITTGAPITVSTYKPLADGGCGGAYDIIICDECHSDTSLIGLITVLDQAEVGARFVVLATATPPGSIPTPHPNIEBVPPLANTGETIPFYAKTIPLEVRGGHLLIPCHSKKKDELPAKLSALGNVAYVYGLDVSVIPASGVVVVYARRGIRYFTPGERSAMPDSVLCEDYAGCANYELTPAETSRLRLAYLNTPGIPVQCDHLEFWSVFTGLTHIDAHPLSOTKQAGDNFYLVAQVATVCARAKAPPSPWDMWKLIRLKPTLHGPTLLYRLGAVONEVTLTHPTIKYIMACMSADLESVTSTWLVGGVIAALAKYCLTGSWVIGRILISGRPAVIDREVLQFDEMEECALPYIEQQNQLAEPOKALGLLQATQKAAAPVSVESKWRALLETFWAKHMNFISGICYLAGSLTLPGNPAISLMAFTSTPLATQYTLFNLGFWAAQLAPPASAAAFVCGAGIAGAAGVSGIHLGVLDILAGYAGVAGALVAFKVMGMDPSTEDLVNLLPALISPGALVGVCAAILRHVGVGEGAVOMNRLIAPAGRGHNVSPHVVVPESDAARVTOILSNLTITQLLKLHWINEDCSTPCSGSLRDVMDWICTVLADFTWQSKLLPRLPFPVPPFSQRTYKRWNRGDIWYTCGQATGHVXNGSMRIVGPRCTSNTHGTPPINAYTTPGCTPTAPNYSRALRWABEYVTRVGDVHYVTGMTDNNKPCQVPAPFETELDGVRLHRYAPACKLDRDEVTFOVLNQYTVGSQLPCEPDPVTVTSMLTDPHSHITAEARRELRGSPPLAGSSASOLSALSKATCTTHGAPDLDLEANLWRQMGNIITRVESNKKIVTLDSPFLRAEDEREVSAAAILKTRKTPAAMPVWARPDYPPLESWKNEDYVPPVVGCPYPPKAPPIPPRRKRTVLTSTVSSALAEATKTGSGSSVDSGTATGPDQAGAGGDAESYSNPPLEGPDPDLSGDSWSTVSEASDHYVCCSWSYTMGALITPCAAEESKIPINALSNPLLRHNMVYSTSRASLROKOTFFDRMVDLDHYRDLVKEMAKASTVKALKSVBEAKLTTPHSKSKPGYKAKOVRSLSRAVNHILRSVMKDLLEDTPIOPTIMAKNEVFCVPEKGRKPARLIVFPDLGVYRCEKMYLDVSTLPOAVMSSYGVFOYSPKQVFEVLPNTWAKKCPMGFSYDTRCFDSTVITENDIRVBESIYQCCDLAPARQIRSLTERLYIIGGPMTNKQNGCYRRCRASGVLTTSQGNLTTCYLKAAAACRAKLODCTMLVCGDDLVICDSAGTQEDASLRVFTAMTRYASPPGDPQPEYDLELITSCSNVSAHDAGKRVYILTRDPTTPPLARAWEHTARHTPVNSMLGNIIMVAPTUWARMILMTHFSLIAQOELEKALDCQIYGATVSIPELDLPQIIQRLHGLSAPLSHSYSPGEINRVASLKLGVPLRLVRWRHARSVRAKLLSQGRRAATCGKYLFWNAVTKLXLITPIPEASQLDLSGNFVAGYSGGDIYHSUSRARPRFMWMLLLLSLVGYILLPNR"		QY 188 TCCTGGACGATCATTTACCGGACGCTGCTCAAGGAGATGAAGCCGAGCGTCCACAGTGA 247	
Db 7774 TCCTGGACGATCATTTACCGGACGCTGCTCAAGGAGATGAAGCCGAGCGTCCACAGTGA 7833		QY 248 AGGCTAAACTGCTATCTGTAGAGAACGATCGAAGCTGACGCCGCCGCGATTCGGGCAAT 307	
Db 7834 AGGCTAAACTGCTATCTGTAGAGAGGATCAAGCTGACGCCGCCGCGATTCGGGCAAT 7893		QY 308 CCAAAATTTGGCTATGGGGCAAGGACGCTCCGGAGCTTATCCAGCAGGCGCGTTAACACA 367	
Db 7894 CCAAAATTTGGCTATGGGGCAAGGACGCTCCGGAGCTTATCCAGCAGGCGCGTTAACACA 7953		QY 368 TCCGCTCGTGTGGAAGACTTGTGTGAGGACACTGACACACCAATTCAGACACCACTCA 427	
Db 7954 TCCGCTCGTGTGGAAGACTTGTGTGAGGACACTGACACACCAATTCAGACACCACTCA 8013		QY 428 TGGCAAAATATGAGGTTTCTGCTGCTCAACACAGAGAAAGGAGCGCGAAACCACTCGCC 487	
Db 8014 TGGCAAAATATGAGGTTTCTGCTGCTCAACACAGAGAAAGGAGCGCGAAACCACTCGCC 8073		QY 488 TCATCGTATTTCCAGACCTGGGAGTTCTGTGTATGCGAGAGATGCGGCTTTTACGACGTG 547	
Db 8074 TTATCGTATTTCCAGACCTGGGAGTTCTGTGTGCGAGAGATGCGGCTTTTACGACGTG 8133		QY 548 TTTCACCTCTTCTCAGCGCTGATGGGCTCTCTCATACGATTCCTACTCTCTTAAGC 607	
Db 8134 TTTCCACTCTTCTCAGCGCTGATGGGCTCTCTCATACGATTCCTACTCTCTTAAGC 8193		QY 608 AGCGGTGAGTTCTGCTGTAATACCTGGAAGCAAAAGAAATGCCCTATGGGCTTCTCAT 667	
Db 8194 AGCGGTGAGTTCTGCTGTAATACCTGGAAGCAAAAGAAATGCCCTATGGGCTTCTCAT 8253		QY 668 ATGACACCCGCTGTTTGAATACCTGGAAGCAAAAGAAATGCCCTATGGGCTTCTCAT 727	
Db 8254 ATGACACCCGCTGTTTGAATACCTGGAAGCAAAAGAAATGCCCTATGGGCTTCTCAT 8313		QY 728 TTTTACCAATGTTGTGACTTGGCCCCCGAGCTAGACAGGCAATAGGTCGCTCACAGAGC 787	
Db 8314 TCTACCAATGTTGTGACTTAGCCCCCGAGCTAGACAGGCAATAGGTCGCTCACAGAGC 8373		QY 788 GGCTTATGTGGGGGTCCCATGACTTAACTCCAAAGGGAGAACTCGCGGTATCGCGGT 847	
Db 8374 GGCTTATGTGGGGGTCCCATGACTTAACTCCAAAGGGAGAACTCGCGGTATCGCGGT 8433		QY 848 GCGCGCGAGCGCGCTGCTGACGATAGCTGGGTAAATACCTTACATGCTACTTGAAGG 907	
Db 8434 GCGCGCGAGCGCGCTGCTGACGACTAGCTGGGCAATACCTTACATGTTACTTGAAGG 8493		QY 908 CCGCTGACGCTGTCGAGCTGCAAGCTCCAGGACTGACAGATGCTGCTGAATGGAGAGC 967	
Db 8494 CCGCTGACGCTGTCGAGCTGCAAGCTCCAGGACTGCAAGTGTGTGTGCGGAGAGC 8553		QY 968 ACCTTGTCTTATCTGTGAAAGCGCGGAAACCAAGAGAGACGCGCAAGCTTACGAGTCT 1027	
Db 8554 ACCTTGTCTTATCTGTGATAGCGGAGACTCAGAGAGCGCGGAGCTTACGAGTCT 8613		QY 1028 TCAGGAGCTATGACTAGGTAATCTTGGCCCCCTTGGGAGACCCGCCAACCGGAATACG 1087	
Db 8614 TCAGGAGCTATGACTAGGTAATCTTGGCCCCCTTGGGAGACCCGCCAACCGGAATACG 8673		QY 1088 ACTTCGAGCTGATACATGCTGTTCTCGAATGTGTGGTTCGACACGATGATCTTGGA 1147	
Db 8674 ACTTCGAGCTGATACATGCTGTTCTCGAATGTGTGGTTCGACACGATGATCTTGGA 8733		QY 1148 AAAGGTGTACTTACCTCACCGTGACCCCGCTTGGCGGGCTGGTGGGAGA 1207	
Db 8734 AACGGGTGTACTTATCTACCGTGACCCCAACCGCTTGGCGGGCTGGTGGGAGA 8793		QY 1208 CAGCTAGGACACACTCCAGTCAACTCTGCTAGGCAACATCATCATGTATGCGGCCACTT 1267	
Db 8794 CAGCTAGGACACACTCCAGTCAACTCTGCTAGGCAACATCATCATGTATGCGGCCACT 8853			

ORIGIN

Query Match		90.9%; Score 1639, 4; DB 14; Length 9548;	
Best Local Similarity		95.2%; Pred. No. 0;	
Matches 1691; Conservative		0; Mismatches 86; Indels 0; Gaps 0;	
QY 8 GCATGTCAATGCTTATACATGACAGCGCCCTGATCACCGTGGCTGCGGAGGAGAA 67		Db 7594 GCTGCTGATGTCCTACATGACAGCGCCCTTATTACCATGCGCGCGAGGAGA 7653	
QY 68 GCAAGCTGCCCATCAAGCGCTGAGCAACTCTTGTGCTGCTGCTCACCATAACCTGGTCTATT 127		Db 7654 GCAAGCTGCCCATTAATGCGCTCAGCAACCTTTGCTGCGCCACACACATGCTATT 7713	
QY 128 CCACACATCCCGCAGTGCAGAGCTCGGAGAGAAAGGTCACTTTTGACAGACTGCAAG 187		Db 7714 CCACACATCCCGCAGTGCAGAGCTCGGAGAGAAAGGTCACTTTTGACAGACTGCAAG 7773	

QY	1268	TGTGGGCAAGGATGATTCTGATGACTCACTTCTTCTCCTCATCTTCTTAGCCCGGAGCAAC	1327
Db	8854	TATGGGCAAGGATGATTCTGATGACCCACCTTCTTCTCCATCTTCTTAGCCCGGAGCAAC	8913
QY	1328	TTGAAAAAGCCCTGGAATGTCAGATCTACGGGGCTGTGTTACTCCATTGAGCACTTGACC	1387
Db	8914	TTGAAAAAGCCCTGGAATGTCAGATCTACGGGGCACTTACTCCATTGAGCACTTGACC	8973
QY	1388	TACCTCAGATCATTGAACGACTCCATGCTCTTAGCGCATTTTCACTCCATGATTACTCTC	1447
Db	8974	TACCTCAGATCATTGAACGACTCCAGCTCTTAGCGCATTTTCACTCCATGATTACTCTC	9033
QY	1448	CAGGTGAGATCAATAGGCTGGCTTCATCCCTCAGAAAGCTTGGGGTACCACCTTGGGAG	1507
Db	9034	CAGGTGAGATCAATAGGCTGGCTTCATCCCTCAGAAAGCTTGGGGTACCACCTTGGGAG	9093
QY	1508	TCTGAGACATCGGGCCAGAGTGTCCGGCTTAAGTTACTGTCCAGGGGGGGAGCGCG	1567
Db	9094	TCTGAGACATCGGGCCAGAGGCTCCGGCTTAAGTTACTGTCCAGGGGGGGAGCGCG	9153
QY	1568	CCATTGTGGCAAGTACCTCTTCAACTGGGAGTAAGGACCAAGCTTAAACTCACTCCAA	1627
Db	9154	CCACTGTGGCAAAATACCTCTTCAACTGGGAGTAAGGACCAAGCTTAAACTCACTCCAA	9213
QY	1628	TTCCGGCTCGCTCCGGCTGGACTTGTCCGGCTGGTTCGTTGCTGGCTACAGCGGGGAG	1687
Db	9214	TCCAGAGGCTCCAGCTGGACTTGTCCGGCTGGTTCGTTGCTGGCTACAGCGGGGAG	9273
QY	1688	ACATATATCACAGCCTGCTCTGTCGGCGACCCCGCTGGTTCATGTTGCTGCTACTCTAC	1747
Db	9274	ACATATATCACAGCCTGCTCTGTCGGCGACCCCGCTGGTTCATGTTGCTGCTACTCTAC	9333
QY	1748	TTCCCGTAGGGTAGGCATCTACTTACTTCCCAACCG	1784
Db	9334	TTTCCCGTAGGGTAGGCATCTACTTACTTCCCAACCG	9370

RESULT 4
AF165059
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AF165059
Hepatitis C virus strain MD8-1 complete genome.
AF165059
AF165059.1 GI:5918956
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus
1 (bases 1 to 9379)
Nagayama,K., Kurosaki,M., Enomoto,N., Maekawa,S.Y., Miyasaka,Y., Tazawa,J., Izumi,N., Marumo,F. and Sato,C.
Time-related changes in full-length hepatitis C virus sequences and hepatitis activity
Virology 263 (1), 244-253 (1999)
20013325
10544098
2 (bases 1 to 9379)
Nagayama,K., Kurosaki,M., Enomoto,N., Maekawa,S., Miyasaka,Y., Sakamoto,N., Fukuma,T., Tazawa,J., Izumi,N., Marumo,F. and Sato,C.
Direct Submission
Submitted (06-JUL-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan
Location/Qualifiers
1..9379
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD8-1"
/db_xref="taxon:111103"

/codon_start=1
/product="polyprotein"
/protein_id="AAD56194.1"
/db_xref="GI:5918957"
translation="MSTNPKPQKTKRNTNRRPQDKVPGGQIIVGGVILLPRRGPRL
GVATKTKSERQPRRRQPIPKARPERGTWAQPGYPMPLVNGEOMGAGMLVSRG
SPVSGPNDPRRSRNLGVIDITLTCGFAFLMGYIPLVGAALGAAALAHAGRVLLER
SVNYATGNLPGCSFSI FLALLSLCTI PASAYEVNRVSGVHTVHTDNCSSIVTAAAD
I I NHTPGCMPCVREDNISRCWALJPTLAARNI SVPTTTIRHVDLLTAATAFCSAMV
VGLDGSFLVLSQLFTFSRRHETVQDCNCSII PGHVSGRHMANMMWNSPTAALV
SOLLRIPOAVDMVAGAHGVLGAYYSMWGNWAKVLIVMLLPAGVDGDTTHVTGTO
AKSAHILTSFLTVGPSQIQLVNTGSMHINRTALNCNDSLOTGFALALFYTHRNAS
GCPERVASCRIPTDPQMGPIITYAEPNNPDORPYCMHYAPRPGCI EPASKVCGPVYC
FTPSVVGTTDRPGVPTWNGENETDVLNANTRPQGNWFGCTWMNSGTGTTKCGG
PPCNI GAGNTLICTPTDCFRKHEPATYTKCSGPHLTPRCULVDVYRLWHVPTCVNF
TIPKVMYVGGVHRLSAACNMWTRGRCDLERSELS PLLLSTTEWQILLPCSFITL
PALSTGLIHLNRINVDVQYLVGVSASFARWEYVLLLLFLLADARVACALMMMLL
IAQABAALENLVINAASVAGAHGILSFLVFFCAAWYIOGRVPGRAYALYGVNPLLL
LLALPRAYAMDRMAASCGGAVFVGLAFLTSLPYKVLARLIMWLQVFTIRAEAH
LQWVPLNVRGGRDALILLMCHVPELIDITKLALIFGLPLMLOAGITRVPYFVR
AHGLIRACMLRVKAGGHVQMAFVKLAALTGTYYVDHLAPLOHMAHSGRLDLAVRE
PVVFSDMETKIIITMGADTAACGDIILGLPVSARRGHEIILGPADSDDEGGWRLLAPIT
AYSQOTRGLIGCIITSLTRDKNOVEGVVSTATQSFATCINGCVMTVFFHAGSK
TLAGPKGPITQMYTNVDQLVGMQAPPGARSMTPTCGSSDLYLIVTRHADVTPVRRR
DGRSLLS PRPVSYLKSGSGPILLCPSGHAGVI FRAAVCTRGVAKAVDFIPVESMET
MRSPFTDINSPPAVQTQVAHLHAPTGSKSTKVPAAVAAQYKVLVLPNSVAATL
GFGAYMSKAYDPENRTGRTITITGAPITITYTGYGLADGGSGGAYDIILCDECHS
TDSITILIGITVDLQDAETAGARLWLATATPGSVTVPHPHNIEVALSNTGEIIPFYK
AIPVIGKGRHILFCHSKKKKDELAALKSLANAVAYRYRLDVSVIPTSGDVVVVGA
TDALMTGTCDFSDVDCNTVTQTVDFSLDPTFTIETTVQDASVRSORRGRTGRG
RGGIYFVTGFRPSGMSVSLCECYDAGCAWELTPTAETS VRLRAYLNTGLPVQC
DHLEFVESVFTGLTHIDAHFLSQTKQAGDNFPVLVAYQATVCARAOA PPPSWQMMKC
LIRLKPITLHGPTPLLYRLGAVNEVLTPTHTTKFMACMSADLEVTSLTWLVGGVLA
ALAAVCLTTGSGVIVGR IILSGRPVITPREVLYREFDEMEECASHLPYIEQGMQLAAE
QFKOKALGLLQATKQAEAAAPVVESSKRALETFAKHMNFISGIQYLAGSLTLPGN
PATASLMAFTASITSLPTTOTILMENTLGGWAAQLAPPASAFAGVAGLAGAAVGI
GAGKVLVDILAGVAGVAGALVAFKMSGEMSPSTEDILNLLPALISGLAVGVVCAA
ILRRHVGPGECAVOMNRLIAFASGRNHVSFTHYVFPESDAAARTQILSLTITOLLK
RLHQWINEDECTPCSGSMLRDVMDICTVLSDFKTLQOSDKIMPLRPGVPLFSLSCQRYK
GWMRGDGMOTTCPCGAQITGHVNGSMRIIVGPKTCSNTWHGTFPFINATYTGPCPTSP
YAPACKPLLEBEVTFVGLNQYLSQLPCEPEPEDVAULTSMLTDPHSHTAETAKRRI
ARGSPPSLASSASSASQLSAPSLKATCTTHDSRSPDADLLEALLWROEGNITRVESEN
KVIILDSFLPRLABEDEREVSAAEILRRSRKFPPLPVPWARPDPNPPLESWKDPDI
ATAPDDQASDDGKESDVESSYMPLEPGFPDPLSDGSWSVTSEASEDWDVCCSM
YTMFGALITPCAAEESKLPIINPLSNLIRHNHNVYATTSRSASLRQKXVTFDRLOQLVI
DHYRDVLKEMKAKASTYKAKLISI BEAKLTTPHARSKFCYGAARDVSLSSKKNALVI
RSVMEDLLEDETETIDITITINAKSEVFCQPEKGRKPARLIVFPLDGVYCEKNALVI
VVSTLPQAVMGSSYQFQSPQRVREFVLTWKSXKCMGFSYDTRCRASGLVTIESDRI
EESIYQCCLDAPAEARKAKISILTERLYIGPLTNSKQNGQYRRRCASGLVTIESDRI
TCYLKSAACRAAKLQDCTMLVNGDLVVICESAGTQEDASLRVFTTEATRYAPAP
DPQPEVDLELIT TSCSSNVVAHDASGRVYVLTTRDPTT PLARAWEFTARHTPVNSWI
GNIIMYAPTLEWMLMTHFFSILLAOBQLEKALDCCOIVGACYSIEPLDLPQI11ERLI
GLSAPLSHSYSPGSEINRVASCRLRGVFPFLRWHRHARSVRAKLISQGGRAATCGKY
FNWAVRTKLKLTPIPAASRLDLSGNFVAGYSGGDIYHLSLRPARPRNFMWCLLLLSVSG
GIYLLPNE"

ORIGIN

Query Match	90.4%	Score	1629.8	DB	14	Length	9379
Best Local Similarity	94.8%	Pred. No.	0				
Matches	1685	Conservative	0	Mismatches	92	Indels	0
				Gaps	0		
QY	8	GCATGTCATGCTTATACATGACAGCGCCCTCATCACACCGTGGCTGGCGAGGAAA	67				
Db	7582	GCTGCTCAATGCTTACACATGACAGCGCCCTCATCACCGCATCGCTGGGAGGAGA	764				
QY	68	GCAAGCTGCCCATCAACGCGCTGAGCAACTCTTCTGCTGCATCAACCATACCTGCTATT	127				

Db 7702 CCACAAATCCCGTAGCGCAAGCCTCGGGCAGAGAAAGGTACCTTTTGACAGACTGCAAG 7761
Qy 188 TCTGTGACGACATTAACCGGAGCGTGTCTCAAGGAGATGAAGGCGAGCGGTCCACAGTGA 247
Db 7762 TCTGTGACGACCACTACCGGAGCGTGTCTCAAGGAGATGAAGGCGAGCGGTCCACAGTGA 7821
Qy 248 AGGCTAAATCGTATCTGTAGAGAGACATGAAGCTGACGCCCGCGCATTCGGCCAAAT 307
Db 7822 AGGCTAAATCTTATCTATAGAGAGAGCGCTGAAGCTGACGCCCGCCCAATTCGGCCAGAT 7881
Qy 308 CCAAAATTTGGCTATGGGCAAGAGCGTCCGGAGCGCTATCCAGCAGGCGCGTTAACCCACA 367
Db 7882 CCAAAATTTGGCTATGGGCAAGAGCGTCCGGAGCGCTATCCAGCAGGCGCGTTAACCCACA 7941
Qy 368 TCGCTCCGTGTGGAAGGACTTGTCTGGAGGACACTGACACACCAATTCAGACCAACATCA 427
Db 7942 TCGCTCCGTGTGGAAGGACTTGTCTGGAGAGACTGAGACACCAATTCAGACCAACATCA 8001
Qy 428 TGGCAAAATAGAGTTTCTCGCTCCAAACAGAGAAAGGAGCGCGCAAAACCGAGCTGCC 487
Db 8002 TGGCAAAATAGAGTTTCTCGCTCCAAACAGAGAAAGGAGCGCGCAAAACCGAGCTGCC 8061
Qy 488 TCATCGTATCCACAGCTGGAGTTCGTGTATGCGAGAGATGCCCTTTACGACGTGG 547
Db 8062 TTATCGTATCCACAGCTGGAGTTCGTGTATGCGAGAGATGCCCTTTACGACGTAG 8121
Qy 548 TTTCCACTCTTCTCAGGCGGTGATGGCTCTCATACGATTCCTCAATCTCTCTAAGC 607
Db 8122 TTTCCACTCTTCTCAGGCGGTGATGGCTCTCATACGATTCCTCAATCTCTCTAAGC 607
Qy 608 AGCGGTCCAGTCTCTCGTGAATACCTGGAAAGCAAGAAATGCCCTATGGCTTCTCAT 667
Db 8182 AGCGGTCCAGTCTCTCGTGAATACCTGGAAAGCAAGAAATGCCCTATGGCTTCTCAT 8241
Qy 668 ATGACACCGGTGTTTGTACTCAACGGTCACTGAGAAATGACATCCGTGTGAGGAGTCAA 727
Db 8242 ATGACACCGGTGTTTGTACTCAACGGTCACTGAGAGTGAATCCGTACTGAGGAGTCAA 8301
Qy 728 TTTACCAATGTTGACTTGGCCCGGAGCTAGACAGGCGCATAGGTCGTCTACAGAGC 787
Db 8302 TTTACCAATGTTGACTTGGCCCGGAGCTAGACAGGCGCATAGGTCGTCTACAGAGC 8361
Qy 788 GGCTCTATGTGGGGGTCCCATGACTACTCAAGGCGCAGAACTGCGGTATCGCGGT 847
Db 8362 GGCTCTATGTGGGGGTCCCATGACTACTCAAGGCGCAGAACTGCGGTATCGCGGT 8421
Qy 848 GCGCGGAGCGCGTGTGACGACTAGCTGCGGTATACCTCAGATGCTACTTTGAAG 907
Db 8422 GCGCGGAGCGCGTGTGACGACTAGCTGCGGTATACCTCAGATGCTACTTTGAAG 8481
Qy 908 CCGTGCAGCGCTGTGAGCTGCCAAGCTCCAGGACTGCAGATGCTCGTGAATGGAGC 967
Db 8482 CCGTGCAGCGCTGTGAGCTGCCAAGCTCCAGGACTGCAGATGCTCGTGAATGGAGC 8541
Qy 968 ACCTGTGCTTATCTGTGAAGCGCGGAGCCCAAGAGGACCGGCAAGCTACGATCT 1027
Db 8542 ACCTGTGCTTATCTGTGAAGCGCGGAGCCCAAGAGGATGCGGCGAGCTACGATCT 8601
Qy 1028 TCAGGAGGCTATCACTAGGTACTCTGCCCGCTGGGACCGCGCCCAACCGGATACG 1087
Db 8602 TCAGGAGGCTATCACTAGGTACTCTGCCCGCTGGGACCGCGCCCAACCGGATACG 8661
Qy 1088 ACTTGGAGCTGATAACATCGTGTCTTCCAATGTGCGTCCGACACAGATGCATCTGGCA 1147
Db 8662 ACTTGGAGCTGATAACATCGTGTCTTCCAATGTGCGTCCGACACAGATGCATCTGGCA 8721
Qy 1148 AAAGGTGTACTACTCACCGGTGACCCCAACCGTCCCGCTGGCGCGGCTCGGTGGAGA 1207
Db 8722 AAAGGTGTACTACTCACCGGTGACCCCAACCGTCCCGCTGGCGCGGCTCGGTGGAGA 8781
Qy 1208 CAGTAGGACACTCCAGTCAATCTGCTGGCTAGGACACATCATGTATGCGGCCACTT 1267
Db 8782 CAGTAGGACACTCCAGTCAATCTGCTGGCTAGGACACATCATGTATGCGGCCACTT 8841

Qy 1268 TGTGGGCAAGGATGATTTCTGATGACTCTCTTCTTCTCCATCTTCTAGCCAGGAGCAAC 1327
Db 8842 TATGGGCAAGGATGATTTCTGATGACTCTCTTCTTCTCCATCTTCTAGCCAGGAGCAAC 8901
Qy 1328 TTGAAAAAGCCCTGGATTGTGAGATCTACGGGGCTTGTACTTCCATTGAGCCACTTGACC 1387
Db 8902 TTGAAAAAGCCCTGGATTGTGAGATCTACGGGGCTTGTACTTCCATTGAGCCACTTGACC 8961
Qy 1388 TACTTCAGATCATTTGAACGACTCCATGCTTTAGCGCATTTTCACTCCATAGTTACTCTC 1447
Db 8962 TACTTCAGATCATTTGAACGACTCCATGCTTTAGCGCATTTTCACTCCATAGTTACTCTC 9021
Qy 1448 CAGGTGAGATCAATAGGGTGGCTTCATGCCCTCAGAGAGCTTGGGGTACCAACCTTGGCAG 1507
Db 9022 CAGGTGAGATCAATAGGGTGGCTTCATGCCCTCAGAGAGCTTGGGGTACCAACCTTGGCAG 9081
Qy 1508 TCTGGAGACATCGGGCCAGAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGGCCG 1567
Db 9082 TCTGGAGACATCGGGCCAGAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGGCCG 9141
Qy 1568 CCATTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAA 1627
Db 9142 CCATTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAA 9201
Qy 1628 TTCGGCTGGCTCCCGCTGGACCTTGTCCGGCTGGTGTGCTGCTACAGCGGGGAG 1687
Db 9202 TTCGGCTGGCTCCCGCTGGACCTTGTCCGGCTGGTGTGCTGCTACAGCGGGGAG 9261
Qy 1688 ACATATATACAGCGCTCTCGTCCGCGAGCCCGCTGGTTCATGTTGCTACTCTCTAC 1747
Db 9262 ACATATATACAGCGCTCTCGTCCGCGAGCCCGCTGGTTCATGTTGCTACTCTCTAC 9321
Qy 1748 TCTCCGTAGGGGTAGGCATCTATCTACTTCCCCCAACCG 1784
Db 9322 TTTCCGTAGGGGTAGGCATCTACTCTCTCCCAACCG 9358

RESULT 5
AF165060
LOCUS AF165060 9379 bp RNA linear VRL 04-NOV-1999
DEFINITION Hepatitis C virus strain MD8-2 complete genome.
ACCESSION AF165060
VERSION AF165060.1 GI:5918958
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1 (bases 1 to 9379)
Nagayama, K., Kuroaki, M., Enomoto, N., Maekawa, S. Y., Miyasaka, Y., Tazawa, J., Izumi, N., Marumo, F. and Sato, C.
Time-related changes in full-length hepatitis C virus sequences and hepatitis activity
Virology 263 (1), 244-253 (1999)
MEDLINE 20013325
PUBMED 10544098
REFERENCE 2 (bases 1 to 9379)
Nagayama, K., Kuroaki, M., Enomoto, N., Maekawa, S., Miyasaka, Y., Sakamoto, N., Fukuma, T., Tazawa, J., Izumi, N., Marumo, F. and Sato, C.
Direct Submission
Submitted (06-JUL-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan
Location/Qualifiers
FEATURES
source
1..9379
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD8-2"
/db_xref="taxon:11103"
/note="from a patient with persistently high serum ALT values; patient 8, point 2 genotype: 1b"

330..9362	/codon_start=1	/product="polyprotein"	/protein_id="AAD56195.1"	/db_xref="GI:5918959"	/translation="MSTNPKPQRKTRNTNRRPQDVKFPGGQIVGGVYLLPRRGPRL GVRAKTSRSRQRRGRQPIKARRPGRTRWAGPGYPMPLVGNBGMWAGMLLSPRG SRSGNDPRRRSRNLGKVIDITLCGFADLMGVIPLVGAPLGAARALAHGVRVLED SVNTATGLPGCSFSIFLLALLSLCTIPASAYEVNRVSVGVHVTNDCSNSIIVTEAD IIMTPGCTPCVCGNISRCWVALTIPASARNISVPTTIRRHVDLVIAAFCSAM VGLDGSFELVSOLFTEPSRHEVQDCNCSIYPCHVSGRHMAWMMNMSPTAALV SQLRIPOAVDMVAGHWGLAGLAYSMVGNKAVLIIVMLPAGVDGTHVTGGAO AHSALHMSFFTVGPSQIKIUVNTNGSHINRTALNCNDSLOTGFLAALFYTHRFNAS GCPERMASCRPIDTFDOGMGPIITAEPSNPDPORPCWYAPRPGCIPASKVGPFVCG FTSPVVGTVTRDFGVPTYNWGENETDVLNNTRPPQGNWPGTWMNSTGFKTCGG PPCNIIGAGNNLTICTPTCFKHPKPEATVTKCGSGPMLTPRCLVDYPRYLWHYPCVNF TIFKRVYGVGVEHFLISAACNWRGERCDLEDRRSELSPLLSTTBMQILPCSFTEL PALSTGLIHLHRNIVDVQYLYGVGSAAVSIARWEYVILLFLLLADARVCACLWMMLL TAQBAALLENLVLNAAAVAGHGLSPVFFCAAWYIKGRVLKGRVAPCAAYALVGVPLLL LILLAPPRAYANDREMAASCGVAVGLAFILTSPPYKVFILARLIWLDQYFPIITRAEH LOVWVPLNVRGRDAIILLMCVWHPBELIDITKLLAIIPGLMVLQAGITRVFYFR AHLIRACMLVKVAGHVQMAFVKLAALGTVYVDHLADLOHWAHSLGRDLAVAVE PVFSDMETKIITWGAADTAAGDIIILGIPVSARRGREILLGPADSLRQGWRLIAPIT AYSQOTRGLGCIITSLTRDNQVGEVQVVSSTATQSFLATCINGVYCVTFVHGAGSK TLAGPKGPIQMYTNVDQDLVGMQAPPGARSMTPCTCGSOLYIVTRHADVYVRRRG DGRSLSLPRPVSYLKGSGGPLLCPSGHVGIFRAAVCTRGVAKAVDFIIPVESMET MSPVFTNSSPPAVPOTFQVAHLHAPTGSGSKTKVPAAYAAQGVKVLVLPVSAEFT GFGAYMSKAYGTDPNIRITGIRITITGAPIYTVSTYKFLADGCGSGAYDIIICECHS TDSITLIGTGLVDAQETAGARLVLAATATPPGSVTVPHNIEBALNNTGCEIIPYK AIPIEVIKGRHLIFCHSKKCDRLAAKLSALGUNAVAYITRGLDVSVLITPSGDVVVVA TDALMTGTGDFSDVIDCNTVOTDFSLDPTFTIETTVPDVAVSRQRRTGRG RGIIVRFVTPGRPSGMDFSSVLCYDAGCANWELTPAETSVRLRAYLNTPGLPVQC DHLEPNESVFTGLTHIDAHFLSOTKQAGNPPYLVAQATVCARAAQAPPSPWDQWKC LIRAPTLHGPTPLVRLGAVQNVTLHPITKTFIMACKSADLEVTWTVLVGVGLA ALAAVCLTGSVVIIVIRIILSKPAVIDREVLYOFDEMECEASHLYPIIEQHQLAB QPKALGILLQATQKAEAAAPVESKWRALETFWKAKMNFISQIYIAGLSTLPGN PAIASLMAFTSITSLPTTQTTMFNILGGMVAQALAPSAASFPVAGIAGAAGVSI GIGKVLVDILAGVAGVAGALVAFKVMGEMPESTEDLINLPAILSPGALVGVVCA ILRRHVQEGGAVQWNRLLIFASRGNHVSPTHVYVESDAAARVTOILSLITOLIK RLHQMINEDCTPCSGWLDRVDWNICTVLIDFKTWLQSKIMPLRPGVPFLSCORGYK GWMRGDMQITPCGAQITGHVKNXGSMRIVGFKTCSTNHWGTFPEINATVDPCTPSP APNYSRALMRVAEEVYEVTRGVDFHYVTGNTNKKPCQVPAPEFFTEGVGVLRH YAPACKLLREVTPOVLNOVLVGSOLPCEPEPDVAVLTSMLTDPESHITAEAKRL ARGSPPLASSASOLSAPLAKATCTTHDSPADLIENALWROEMGNTIRVSEEN KVILDSFPLRAEDEREVSAAELLRSRKEFPALPVWARPDPYNPILLESKWDPDY VPVPHGCPPLTPKAPITPPRRKRTVLTESVSSALAEALATKTFSGSGSADVDSGT ATAPPQASDGDGKESDYESYMPLEGEQDPLSDGMSWTSEESASEDVCVCSMS YWTGALITPCAAEASKLPINLSNLLRHNMYATTSRSASLRKKKVTDFRLQVLD DHYRDLKEMKAKASTVKAKLLIREACKLTPHGSARKFYGAKDVRSLSKAVNHI RSVWEDLLETETPIDTTIMAKSEVPCVOPEKGRKPARLIVFPDLGVRVCEKMDIYD VYSTLPQAVNGSFOYSPKQRFELVNTWKSXKCPMGFSYDTRCFDSTVTESDIRT BESIYQCDLAPAEARKAIKSLTERLYIGGLYKNSQNGCYRRCRASGVHLTSCNTLI TCVQLATAACRAKLQDCTMLVNGDLVVICESAGTQEDASLRVPTAMTRYSAAPP DPQPPVDLELITSCSNVSVAHDAGKRVYLTTRDPTPLARAAWETARHTHPVNSWL GNLIWAPTLWARMILMTHFSLILAOEOLEKALDCOYIGACVSEIPLDLPOLIEBRLH GLSFAFLSHYSYSGEINRVASCLRLKLVPLRVWRHARSVRAKILSVQGRAATCKYL FNWAVRTKGLITPIPAASRLSDLGWFMFVAGTSGGDIYHLSLRARPRWFMKLLLSVGV GIVLLEPNR"	Db	7702	CCACAACATCCCGTAGCGCAAGCTCGCGGAGAAAGGTCACCTTTGACAGACTGCAAG	7761
Qy	188	TCTTGCAAGATCATTACCGGGACGTGCTCAAGAGATGAAGCGGAGCGTCCACAGTCA	247	7821	7821				
Db	7762	TCTTGCAAGACCACTACCGGGACGTGCTCAAGAGATGAAGCGGAGCGTCCACAGTCA	7821	7821	7821				
Qy	248	AGGCTAAACTGCTATCTGTAGAAAGACATGCAAGCTGACGCCCCCGCATTCGGCCAAAT	307	7881	7881				
Db	7822	AGGCTAAACTCTATCTATATAGAAAGACCTGCAAGCTGACGCCCCCACATTCGGCCAGAT	7881	7881	7881				
Qy	308	CCAAATTTGGCTATGGGGCAAGGACGTCCGGAGCCTATCCAGCAGGGCCGTTAAACACA	367	7941	7941				
Db	7882	CCAAATTTGGCTATGGGGCAAGGACGTCCGGAGCCTATCCAGCAGGGCCGTTCAACACA	7941	7941	7941				
Qy	368	TCCGCTCCGTGTGGAAGGACTTCTCTGAGGACACTGACACACCAATTCAGACCACCATCA	427	8001	8001				
Db	7942	TCCGCTCCGTGTGGAAGGACTTCTCTGAGGACACTGACACACCAATTCAGACCACCATCA	427	8001	8001				
Qy	428	TGGCAAAAATGAGGTTTCTCGCTCCCAACACAGAAAGAGGCGCGCAACACAGCTCGCC	487	8061	8061				
Db	8002	TGGCAAAAATGAGGTTTCTCGCTCCCAACACAGAAAGAGGCGCGCGCAACACAGCTCGCC	487	8061	8061				
Qy	488	TCATCGTATTCCCAAGACCTCGGAGTTCTGTATCCGAGAAAGATGGCCCTTTACGAGCTGG	547	8121	8121				
Db	8062	TTATCGTATTCCCAAGACCTCGGAGTTCTGTATCCGAGAAAGATGGCCCTTTACGAGCTGG	547	8121	8121				
Qy	548	TTTCCACTTCTCCTCAGCGCGTGAATGGGCTCTCATACCGAATTCGAATATCTCTCCTAAGC	607	8181	8181				
Db	8122	TCTCCACCTTCTCCTCAGCGCGTGAATGGGCTCTCATACCGAATTCGAATATCTCTCCTAAGC	607	8181	8181				
Qy	608	AGCGGCTCGAGTTCTCTCGTGAATACCTGGAAGAAAGAAATCCCTATCGGCTTCTCAT	667	8241	8241				
Db	8182	AGCGGCTCGAGTTCTCTCGTGAATACCTGGAAGAAAGAAATCCCTATCGGCTTCTCAT	667	8241	8241				
Qy	668	ATGACACCCCTGTTTGTGACTCAACGGTCACTGAGAATGACATCCGCTGTTGAGGAGTCAA	727	8301	8301				
Db	8242	ATGACACCCCTGTTTGTGACTCAACGGTCACTGAGAATGACATCCGCTGTTGAGGAGTCAA	727	8301	8301				
Qy	728	TTTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCCATTAAGTTCGCTACAGAGC	787	8361	8361				
Db	8302	TTTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCCATTAAGTTCGCTACAGAGC	787	8361	8361				
Qy	788	GGCTCTATGTCGGGGTCCCAATGACTAACTCCAAAGGGGAGAACTCGGCTATCGCCGCT	847	8421	8421				
Db	8362	GGCTCTATGTCGGGGTCCCAATGACTAACTCCAAAGGGGAGAACTCGGCTATCGCCGCT	847	8421	8421				
Qy	848	CGCGCGGAGCGCGTCTGACGACTAGCTGCGGTAAATACCTCACAATGCTTACTTTGAAG	907	8481	8481				
Db	8422	CGCGCGGAGCGCGTCTGACGACTAGCTGCGGTAAATACCTCACAATGCTTACTTTGAAG	907	8481	8481				
Qy	908	CCGCTGACGCTGTCGAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAATCGAGACG	967	8541	8541				
Db	8482	CCACTGCGGCTGTCGAGCTCGGAAGCTCCAGGACTGCACGATGCTCGTGAATCGAGACG	967	8541	8541				
Qy	968	ACCTTGTGCTTATCTGTGAAAGCGGGAAACCCAGAGGAGCGGGCAAGCCTACAGTCT	102	8601	8601				
Db	8542	ACCTTGTGCTTATCTGTGAAAGCGGGAAACCCAGAGGAGCGGGCAAGCCTACAGTCT	102	8601	8601				
Qy	1028	TCACGAGGCTATGACTAGGTACTTCTGCCCTTGGGGACCCGCCCAACCGGAATACG	108	8661	8661				
Db	8602	TCACGAGGCTATGACTAGGTACTTCTGCCCTTGGGGACCCGCCCAACCGGAATACG	108	8661	8661				
Qy	1088	ACTTGGAGCTGATACATCGTGTCTCCCAATGCTCGGTGCGCACACGATGCATCTGGCA	114	8721	8721				
Db	8662	ACTTGGAGCTGATACATCGTGTCTCCCAATGCTCGGTGCGCACACGATGCATCTGGCA	114	8721	8721				
Qy	1148	AAAGGCTGTACTTACCTCACCCCGTGCACCGCTCCCTTTCGCGGGCTCGTGGGAGA	120	8781	8781				
Db	8722	AAAGGCTGTACTTACCTCACCCCGTGCACCGCTCCCTTTCGCGGGCTCGTGGGAGA	120	8781	8781				
Qy	1208	CAGCTAGGCACACTCCAGTCAACTCTCTGCTAGGCAACATCATCATGTATGCGCCACTT	126	8811	8811				

ORIGIN

Query Match

Best Local Similarity 90.3%; Score 1628.2; DB 14; Length 9379;

Matches 1684; Conservative 94.8%; Pred. No. 0;

Mismatches 93; Indels 0; Gaps 0;

Qy	8	GCATGTCATGTCCTATACATGACGAGCGGCGCCCTGATCACACCGTGGCTCGCGAGGAAA	67
Db	7582	GCTGCTCAATGTCCTACATGACGACGAGCGGCGCCCTGATACGCCATGCGCTGGGAGGAAA	7641
Qy	68	GCAAGCTGCCCATCAACCGCGCTGAGCAAATCTCTTGTGCGGTCCACCATAACTGGTCTATT	127
Db	7642	GCAAGTGGCCCATCAACCGCTTGAGCAAATCTCTTGTGCGGTCCACCAACATGGTCTACG	7701
Qy	128	CCACAACATCCCGAGTGCAAGCTCGCGGAGAAAGGTCACCTTTGACAGACTGCAAG	187

```
Db 8782 CAGCTAGACACACTCTCGCTCACTCACTCTCGCTAGGCAACATCATCATGTATGCGCCACCC 8841
Qy 1268 TGTGGCAAGATGATTCGTGATGACTCACTTCTTCTTCCATCTCTTAGCCAGGACCAAC 1327
Db 8842 TATGGCAAGATGATTCGTGATGACTCACTTCTTCTTCCATCTCTTAGCCAGGACCAAC 8901
Qy 1328 TTGAAAGACCCCTGGATTCGTGATGACTCACTTCTTCTTCCATCTCTTAGCCAGGACCAAC 1387
Db 8902 TTGAAAGACCCCTGGATTCGTGATGACTCACTTCTTCTTCCATCTCTTAGCCAGGACCAAC 8961
Qy 1388 TACCTCAGATCATTTGAACGACTTCCATGCTCTTAGGCGCATTTTCACTTCCATAGTTACTCTC 1447
Db 8962 TACCTCAGATCATTTGAACGACTTCCATGCTCTTAGGCGCATTTTCACTTCCATAGTTACTCTC 9021
Qy 1448 CAGGTGAGATCAATAGGCTGGCTTCACTGCTCAGGAAGCTTGGGCTACACCCCTTGGCAG 1507
Db 9022 CAGGTGAGATCAATAGGCTGGCTTCACTGCTCAGGAAGCTTGGGCTTCAACCTTGGCAG 9081
Qy 1508 TCTGGAGACATCGGCGCAGAAAGTGTCCGGCTAAGTTACTGTCCAGGGGGGAGGGCG 1567
Db 9082 TCTGGAGACATCGGCGCAGAAAGTGTCCGGCTAAGTTACTGTCCAGGGGGGAGGGCG 9141
Qy 1568 CCAATTGTGGCAAGTACCTTCACTCGGCGAGTAAGCAACCAAGCTTAACTCACTCCAA 1627
Db 9142 CCAATGTGGCAAGTACCTTCACTCGGCGAGTAAGCAACCAAGCTTAACTCACTCCAA 9201
Qy 1628 TTCCGGCTGCTCCGGCTGGACTTGTCCGGCTGGTGTCTGCTGCTACAGCGGGGAG 1687
Db 9202 TCCCGCTGCTCCGGCTGGACTTGTCCGGCTGGTGTCTGCTGCTACAGCGGGGAG 9261
Qy 1688 ACATATATACAGCTGTCTGCTGCTCCGACCCCGCTGCTTCACTTGTGCTTACTTCTTAC 1747
Db 9262 ACATATATACAGCTGTCTGCTGCTCCGACCCCGCTGCTTCACTTGTGCTTACTTCTTAC 9321
Qy 1748 TCTCGTAGGGTAGGATCTATCTACTTCTCCCAACCG 1784
Db 9322 TTTCCGTAGGGTAGGATCTACTCTCTCCCAACCG 9358

RESULT 6
AB057599
LOCUS Hepatitis C virus gene for RNA-dependent RNA polymerase, partial
DEFINITION cds, clone:pt. 2.
AB057599
VERSION AB057599.1 GI:29466654
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Watanabe,K. and Yoshioka,K.
TITLE Mutation in HCV NS5B
JOURNAL Unpublished
REFERENCE
AUTHORS Watanabe,K., Yoshioka,K., Yano,M., Ukai,K. and Ito,H.
TITLE Direct Submission
JOURNAL Submitted (16-MAR-2001) Kazumasa Watanabe, Nagoya University School
of Medicine, Third Department of Internal Medicine; 65 Tsuruma-cho,
Showa-ku,Nagoya, Aichi 466-8550, Japan
(E-mail:kankenmed.nagoya-u.ac.jp, Tel:81-052-7442190,
Fax:81-052-7442208)
FEATURES
source
1..1773
/organism="Hepatitis C virus"
/virion
/mol_type="genomic RNA"
/db_xref="taxon:11103"
/clone="pt. 2"
/country="Japan"
/notes="genotype 1b"
1..1773
gene
```

```

/gene="RdRp"
<1..>1773
/gene="RdRp"
/codon_start=1
/product="RNA-dependent RNA polymerase"
/protein_id="BAC66793.1"
/db_xref="GI:29466655"
/translaton="SMSYTWTTGALI.TPCAAEESKLPIINALSNSLLRHNLVYTTTSRS
AGLRQKVTEDRLQVLDHVDVLEKMKAKASTVKARLLSVEEACKLTPPHSAKSKFG
YGLDVRNLSKATNHIRSVWMDLLESDPTIDTTIMAKNEVFCVQPEKGRKPARLI
VFDLGVVCEKMAIYDVSTLPOAVMGSSYFQYSPQRVEFLVNTWKAKCPMGFS
YDFCFSTVTENDIRVEESTYQSCDLAPARQAIRSLTERLYVGGPLTNKSGONCGY
RRCRASGLVITSCGNLTCTVYKASACRAAKLQDCTMLVNCDDLAVLICESAGBEDAA
SLRVFTAMRYSPAGDLPOPEYDLELITSCSNVSAVHDAKRVITYLTRDPTPL
ARAWEATARTHTPVNSWLGNIIMAPTUWARMILMTHFFSLILAQOELEKADCOIYGA
YSTEPLDLPOILERLHLSAFSLHSYSGEINRVASCLKGLGYPPIRVVHRARSVR
AKLSQGGRAATCGKYLFWAVRTKLKLTPIPAASRLDLSGFWAGVSGGD1YHLSLR
ARPRWFMLCLLLLSVGVGIYLLPNR"

ORIGIN
Query Match 90.3%; Score 1628; DB 14; Length 1773;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1682; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 13 TCAATGTCCTATACATGACGACGCGCCCTGATCACCGTGCCTGCGGAGGAAGCAAG 72
Db 1 TCAATGTCCTATACATGACGACGCGCCCTGATCACCGTGCCTGCGGAGGAGCAAG 60
Qy 73 CTGCCCATCAACGCGCTGAGCAACTCTCTGCTGCCTCACCACCACTTGTCTATTTCACA 132
Db 61 CTGCCCATCAACGCGCTGAGCAACTCTTCTGCTGCCTCACCACCACTTGTCTATTACA 120
Qy 133 ACATCCCGAGTCAAGCCCTGCGGCAAGAGTCACTTTTGACAGACTGCAAGTCTG 192
Db 121 ACATCCCGAGTCAAGCCCTGCGGCAAGAGTCACTTTTGACAGACTGCAAGTCTG 180
Qy 193 GACGATCACTACCGGAGTGTCTCAAGGAGATGAAGCGAAGCGTCCACAGTGAAGGCT 252
Db 181 GACGATCACTACCGGAGTGTCTCAAGGAGATGAAGCGAAGCGTCCACAGTGAAGGCT 240
Qy 253 AAATCTGCTATCTAGAGAGATGCAAGCTCAAGGAGTCAAGGAGTCAAGGAGTCAAGGAGT 312
Db 241 AGATCTGCTATCTAGAGAGATGCAAGCTCAAGGAGTCAAGGAGTCAAGGAGTCAAGGAGT 300
Qy 313 TTTGGCTATGGGCAAGGAGTCCGAGGCTTATCCAGAGGCGCGTAAACCATATCCGC 372
Db 301 TTTGGCTATGGGCAAGGAGTCCGAGGCTTATCCAGAGGCGCGTAAACCATATCCGC 360
Qy 373 TCCGTGTGAAGGACTTGTCTGAGGAGTCTGACACACCAATTCAGACCACTCATGGCA 432
Db 361 TCCGTGTGAAGGACTTGTCTGAGGAGTCTGACACACCAATTCAGACCACTCATGGCA 420
Qy 433 AAAAATGAGGTTTCTGCTTCCACAGAGAGGAGGCGCGCAACAGCTGCGCTCATC 492
Db 421 AAGAATGAGGTTTCTGCTTCCACAGAGAGGAGGCGCGCAACAGCTGCGCTCATC 480
Qy 493 GTATTCCAGACTGCGGAGTTCGTGTATCGGAGAGATGCGCCCTTTACAGCTGCTTCC 552
Db 481 GTATTCCAGACTGCGGAGTTCGTGTATCGGAGAGATGCGCCCTTTACAGCTGCTTCC 540
Qy 553 ACTCTTCTCTCAGGCGCTGATGGGCTCTCATACGGATTCGAATACCTCTTCTTAAGCAGCG 612
Db 541 ACCCTTCTCTCAGGCGCTGATGGGCTCTCATACGGGTTCCAGTACTCTCTTGGAGCGG 600
Qy 613 GTCGAGTTCCTCGTGAATACCTGGAAAGCAAGAAATGCGCTATGGGCTTCTCATATGAC 672
Db 601 GTCGAGTTCCTCGTGAATACCTGGAAAGCAAGAAATGCGCTATGGGCTTCTCATATGAC 660
Qy 673 ACCCGCTGTTTGACTCAACGCTCACTGAGATGACATCCGCTGTTGAGAGTCAATTTAC 732
Db 661 ACCCGCTGTTTGACTCAACGCTCACTGAGATGACATCCGCTGTTGAGAGTCAATTTAC 720
Qy 733 CAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCCAATAAGTCTGCTACAGAGCGGCTC 792
```


Db 721 CAATCTTGACTTGGCCCCGAAAGCCAGACAGGCTTAAGGTCGCTCACGGAGCGGCTT 780
Qy 793 TATGTCGGGGTCCCATGACTAACTCCAAAGGGCAGAACTCGGCTATCGCGGTGCGC 852
Db 781 TATGTCGGGGTCCCTTAACATAATCCAAAGGGCAGAACTCGGCTACCGCGGTGCGC 840
Qy 853 GCGACGCGGTGTCGACGACTAGCTGCGGTAATACCTTCACATGCTACTTGAAGCGCGCT 912
Db 841 GCGACGCGGTGTCGACGACTAGCTGCGGTAATACCTTCACATGCTACTTGAAGCGCTCT 900
Qy 913 GCGACGCGGTGTCGACGACTAGCTGCGGTAATACCTTCACATGCTACTTGAAGCGCGCT 972
Db 901 GCGGCTGTCGAGCGCGGAACTCCAGGACTGCAATGCTCGTGAAGCGGAGGAGCTT 960
Qy 973 GTCGTTATCTGTGAAGCGCGGGAACCCAGAGGCGGGCAAGCTTACGAGTCTTCACG 1032
Db 961 GTCGTTATCTGTGAAGCGCGGGAACCCAGAGGAGTGGGCGAGCTTACGAGTCTTCACG 1020
Qy 1033 GAGGCTATGACTAGTACTGCTGCCCCCTCGGGACCGCCCCAACCGGAATACGACTTG 1092
Db 1021 GAGGCTATGACTAGTACTGCTGCCCCCTCGGGACCTGCCCCAACCGGAATACGACTTG 1080
Qy 1093 GAGCTGATAACATCTGTTCTTCCAAATGTCGCGTGCACACGATGCACTCTGGCAAAAG 1152
Db 1081 GAGCTGATACATCATGTTCTTCCAAATGTCGCGTGCACACGATGCACTCTGGCAAAAG 1140
Qy 1153 GTGTACTACTCACCGGTGACCCCAACCGTCCCTTTGCGGCGGCTGCTGGGAGACAGCT 1212
Db 1141 GTGTACTACTCACCGGTGACCCCAACCGTCCCTTTGCGGCGGCTGCTGGGAGACAGCT 1200
Qy 1213 AGGCACTTCCAGTCAACTCTCGGTAGGCAACATCATGATGATGCGGCCACTTTGG 1272
Db 1201 AGACACACTCCAGTCAACTCTCGGTAGGCAATATAATCATGATGATGCGGCCACTTTGG 1260
Qy 1273 GCAAGGATGATCTGATGACTCACTTCTTCCATCTTCTAGCCACGAGCAACTTCAA 1332
Db 1261 GCGAGAATGATCTGATGACTCACTTCTTCCATCTTCTAGCCACGAGCAACTTCAA 1320
Qy 1333 AAAGCCCTCGATGTCAGATCTACGGGGCTTGTTACTTCCATTTAGGCCACTTGACCTACT 1392
Db 1321 AAAGCCCTCGATGTCAGATCTACGGGGCTTACTTCCATTTAGGCCACTTGACCTACT 1380
Qy 1393 CAGATCAITGAACGACTCCATGCTTACGGGATTTTCACTCCATGTTACTTCTCCAGT 1452
Db 1381 CAGATCAITGAGCGACTCCATGCTTACGGGATTTTCACTCCATGTTACTTCTCCAGT 1440
Qy 1453 GAGATCAATAGGTGCTTATGCTTACGGGCTCAGGAGCTTGGGTACCCCTTGGAGTCTGG 1512
Db 1441 GAGATCAATAGGTGCTTATGCTTACGGGCTCAGGAGCTTGGGTACCCCTTGGAGTCTGG 1500
Qy 1513 AGACATCGGCGCAGAAAGTGTCCGCGCTAAGTTACTTGTCCACGGGGGAGGCGCCACTT 1572
Db 1501 AGACATCGGCGCAGAAAGTGTCCGCGCTAAGTTACTTGTCCACGGGGGAGGCGCTGCT 1560
Qy 1573 TGTGGCAAGTACCTTCTCAACTCGGCGAGTAAGCAACCAAGCTTAACTCACTCCAATTCG 1632
Db 1561 TGTGGCAAGTACCTTCTCAACTCGGCGAGTAAGCAACCAAGCTTAACTCACTCCAATTCG 1620
Qy 1633 GCTCGGCTCCGGCTGAGTACTTGTCCGGCTGGTTCGTTGCTGCTGCTACAGCGGGGAGCA 1692
Db 1621 GCGGCGTCCGGTGGACTTGTCCGGCTGGTTCGTTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Qy 1693 TATCAGACCTGCTGCTGCTCCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1752
Db 1681 TATCAGACCTGCTGCTGCTCCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Qy 1753 GTAGGGGTAGGATCTATCTACTTCTCCCAACCG 1784
Db 1741 GTAGGGGTAGGATCTATCTACTTCTCCCAACCG 1772

AB049088
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
PUBMED
AUTHORS
TITLE
JOURNAL

FEATURES
source

5'UTR

CDS

AB049088 9616 bp RNA linear VRL 22-AUG-2002
Hepatitis C virus genomic RNA, complete genome, isolate:HCVT094.
AB049088 GI:11559442
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K.,
Hatahara, I., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and
Mishiro, S.
Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
with hepatocellular carcinoma: the 'progression score' revisited
Hepatol. Res. 20 (2), 161-171 (2001)
1134851
2 (bases 1 to 9616)
Mishiro, S.
Direct Submission
Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,
Department of Medical Sciences; 6-3-22 Higashi Oh-i, Shinagawa-ku,
Tokyo 140-8522, Japan (E-mail:shunji.mishiro@po.toshiba.co.jp,
Tel:81-3-3764-9881, Fax:81-3-3764-8992)
Location/Qualifiers
1. .9616
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="HCVT094"
/specific host="human"
/db_xref="taxon:11103"
/tissue type="serum"
/note="HCV-genotype: 1b"
1. .341
/note="complete"
342. .9374
/codon_start=1
/product="polyprotein"
/protein_id="BAB18801.1"
/db_xref="GI:11559443"
/translation="MSTNPXPQRKTKRNTNRRPDVVFPGCGQIVGGVYLLPRGPRI
GVTRTKTSRSQPRRRQIPKAREPEGTWAGPYWLYGNEGMAWGLLSLSPRC
SRPSWGTDPRRSRNLGKVIDTLCTCFADLMGYIPLVGAFLGCAALAHAGVRLV
SNVATCNLPCCSFIFLLALLSCLTTPASAYEVHNVSGIYHVTNDCSNASIVVEAAI
MIMHAPGCPVREDNARSCWALPTLAARNASVPTTIRRHVDLAVGAAAPCSAMV
VGDLCGSVFLISQLFTFSPRRHETVQDCNCSIVPGHVTGHRMANDNMNSPTTALV
SOLLRIPOAVMDMVAGAHGVLGALAYYSMAGNWKVLI VMLLPAGVDGNTVSGAVC
GRTVSSLANIFSLGASQKIQLINTNGSHNIRNTALNCNDSLQTFGLAALFYAHKFN
GCPERMASCRSIDKFDGNGPITVAQPSNLDQRPQWMAQQQCIIPASQCGPVPV
FTSPVVGTTDSSGVPYTSWGENETDVLNNTRPPQGNWFGCTWNSNFTGFKTCQ
PPCNIIGVGNLTLCPTDCFRKHEATYTKGSGSPMLTPRCLVDYPVRLMHPVCTVNI
TIFKRYVGGVHRHAACNWTGRCDLREDRSELSPLLLSITTEWQILPCSFTHI
PALSTGLIHLQNTVDQVLYIGISAVSFAIKWEYVLLPLLDLADALVACLMWMLI
IAQEAALENLVLAASVAGAHGFLVFPFCAAWYIKGRLVPGVAYALVGMWPLLI
LILLALPRAYAMDREMAASCGGAVFVGLAILTSPYKVLARLIWMLOVFIARAEN
LQWVPFLNVRGGRDAIILLTCAAHPELIIDI TKLILAILGLPMVFOAGIIRVPVFI
AQGLIRVCLMVRKVGAGHYVQMLALALTGTVYDHLFLRDNAAHAGLRDLAVAT
PVFSDMETKIIITWAGDTAACDIIISGLPVSARRGKETILGSPADSLSEGRHRLIAVI
AVSQDTEKLCIITSLTRDKNQVEGEVVSATQSFLLASVNGVCTVTVYHAGS
TLAPKGPITOMYTNVDQLVGOAPPGARSMPTCTCCSSDLIVLTHRADYIPVRRR
DGRSLLSPRPVSVLYKSSGGPLCPSHVVGII FRAAVCTRGVAKAVDFVVESEMT
MRSVFTDNSSPPAVPQTFQVHLHAPGSGCKTKVPAVAAQGVKVLNPSVAAIT
GFGYMSKAHGTDPNIITGIRITITGAPITTYTCKFLADGCGCGGAYDIIICDECH
TDSITLIGITVLDQATAGARVYLATATPGSVTPHPNIEVALSNTGEIIPFG
AIPTEAIKGGRHLPCHSKKCDLAKSLGLINAIAYRGLDVSIVTSGDVVVVV
TDALMTGTGPDSDVIDCNTCTVOTDPSLDPTFTIETTVDPQDAVSRTQRRGTGR
ARGIYRVPYPERSGMSFDSVLCEDYDAGCANVELTPAETSVRLLRVLNTPGLPVC
DHLSEWESVFTGLTHIDAHFLSQTQKQDNFPYLVAQATVYCARAAQAPPSDDQWVK
LIRLPTLHGPTPLDLRGAVNEVTLTHPVTKYIMACMSADLEVTSTWLVGLV
ALAAVCLTGSWIVIGRIILSGRPVILPDRVLYREFDEMBECASHLPYIQGMOLA
QPKOKALGLLOTATTQAEAAAPVVESEKWRFALEAFWAKHWNFI SGIQIYGLGLSTLP
PAIASIMAFASITSPITTSQTLFNILGGWVAAQLAPPASAAPVAGIAGAAVGS

QY 1688 ACATATATCACAGCTGCTCTGTCGCCGACCCCGCTGTTTCATGTTGTGCTACTCTCTAC 1747
|||||
Db 9274 ACATATATCACAGCTGCTCTGTCGCCGACCCCGCTGTTTCATGTTGTGCTACTCTCTAC 9333
|||||
QY 1748 TCTCGTAGGGGTAGGCATCTATCTACTCTCCCAACCGG 1785
|||||
Db 9334 TTTCTGATAGGGGTAGGCATCTACTCTCTCCCAACCGG 9371
|||||
RESULT 8
AF207763 9377 bp RNA linear VRL 27-APR-2000
LOCUS Hepatitis C virus strain MD22 complete genome.
DEFINITION AF207763
ACCESSION AF207763
VERSION AF207763.1 GI:7650243
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (bases 1 to 9377)
AUTHORS Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Marumo,F. and Sato,C.
Characteristics of hepatitis C viral genome associated with disease progression
Unpublished
REFERENCE 2 (bases 1 to 9377)
AUTHORS Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Izumi,N. and Sato,C.
Direct Submission
Submitted (23-NOV-1999) Second Department of Internal Medicine,
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-Ku,
Tokyo 113-8519, Japan
FEATURES
Location/Qualifiers
1..9377
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD22"
/db_xref="taxon:11103"
/note="isolated from a patient with liver cirrhosis and hepatocellular carcinoma (HCC-2)"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="AAF65953.1"
/db_xref="GI:7650244"
/translation="MSTNPKPQRKTKNTRRPQDKPFGGQIYGGVYLPRGPHL
GVRAKTSERSQPRQRPPIKARQPEGRAWAQPGYWPPLYGNGMGWAGMLLSPRG
SRPSPNDPRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVED
GVNATGNLPGCSFSIFLLALLSLTTPASAVEENVSGIYHVTNDCSNTSIVYEAD
VINHIHPCVPCVRENNISRWVLAFTLAARNISVPTTMMRHVDLLVCAAFCSAMY
VGLDGSFLLISQLFSTFSRPHETVODCNLSYPGHVSGHMAWMNWSPTTALVL
SOLLRIPOAVDNDVAGAHGVLGAIYISMVGNWAKVLLVLLPAGVDADTHVTGYY
AHFTQGFALFAPGASQNLINTNGSMHNRINALNCNLSLKTGLFUALFYTHFNAS
GCPERMASCRPIQDFQGWGPIYITBPNSPDQRPYCMWYAPRPGICVPASGVPVYC
FTSPVVVGTDSRSGVPTYSWGENETDVLNNTRPPQGNWFGCTWMNSTGFTKCGG
PPCNIGAGNNTLLICTDCEFRKHPEATYTKCGSPWLTPRCLVDVYFRLMHPYCNVF
TIFKVRMYVGVGVEHLRNLAACTNTRGERCDLEDRSELSPLLLSTEWQILPCEFTLL
PALSTGIHLHNRIVDVQYLYGIGSAVSFAIKWEYVYVLLFLLDADARVACLMMMLL
IAQAEALENLVVINAASVAGHGTLSFLVFFCAAWYIKGRVLPVGAAYAFYGVWPLL
LLALPRAYAMDREMAAASCGAVFVGLALLTILSPYKVFARLLIWLQYFTRAEAH
LOWMIPLVNVRGGRDAIILITCAVHPELFDITKLLAILGLPLQWHAHAGRLDTVAPE
AQLGIRACMLVRQAAGHYVOMALVKLAALTGTYYVDHLTPHQWHAHAGRLDTVAPE
PVYFSDMETKIITWGDITACGDIIILGLPVSAARRGRIILLPADSLEGGWRLLPAT
AVSQQTRGLLGCIITSLTGRDNQVGEQVWSTATOSFLATCTVNGVCWTVYHGAASK
TLAGPKGPITMYTNVDQILVWQAPFGARSMPTCTCGSSDLYLVTRHADVIPVRRRT
DGRGSLSPREVSYLKSGSGPLCLPLGHVWGI FRAAVCTRGVAKAVDFVPVSEMT
MRSPVFDNSSPAPVPTQVVAHLHAPTKGSKSTKVPAAVAAQYKVLNPSVAATL
GFGAYMSKAYGTDPNIRITGRTITTCAPITYSTYKFLADGGSGGAYDIIICDECHS
TDSITILIGITVDLQAEATAGARLVLATATPPGGSVTFVPHNIEVALSTTGELPFVYGK

ORIGIN

Query Match 89.8%; Score 1618.6; DB 14; Length 9377;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 1678; Conservative 0; Mismatches 99; Indels 0; Gaps 0,
QY 8 GCATGTCATATGCTCTATACATGGACGGCGCCCTGATCACACCGTGGCGTGGCGAGGAAA 67
Db 7582 GCTGCTCCATGTCATACACATGGACGGCGCCCTGATCACGCCATGCGCTGGGAGAGAGA 7641
QY 68 GCAAGTGGCCCATCAACGCGCTGAGCAATCTCTTGTGCGTCCACCAATCAACTGCTCTATT 127
Db 7642 GCAAGTTGCCCATCAACGCGTTGAGCAATCTCTTGTGCTGCGTCAACCAATATATGCTCTACT 7701
QY 128 CCACAAATATCCCGAGTGCAGAGCTCGCGGAGAGAAAGAGTTCACCTTTGACAGACTGCAAG 187
Db 7702 CCACAAATATCCCGTAGCGCGGCGCTCGCGGAGAGAAAGAGTTCACCTTTGACAGACTGCAAG 7761
QY 188 TCCTGGAGCATCATATACCGGACGCTGCTCAAGGAGATGAAGGGGAGGCGGTCCACAGTGA 247
Db 7762 TCCTGGAGCATCATATACCGGACGCTGCTCAAGGAGATGAAGGGGAGGCGGTCCACAGTGA 7821
QY 248 AGGCTAACTGCTATCTGTAGAAGAAGTCAAGTCAAGTCCGCCCGCCGATTCGGCGCAAT 307
Db 7822 AGGCTAACTCTCTATCTGTAGAAGAAGTCTCAAACTGACGCCGCCACATTCGGCGCAT 7881
QY 308 CCAAAATTTGGCTATGTTGGGCAAGGACGCTCCGAGCGCTATCCAGCAGGCGCTTTAAACACA 367
Db 7882 CCAAAATTTGGCTATGTTGGGCAAGGACGCTCCGAGCGCTGTCAGCAGGCGCGTCAACACA 794
QY 368 TCCGCTCCGCTGGAGAGACTTGTGTGAGGACACTGTGACACACCAATATTCAGACCACTCA 427
Db 7942 TCCGCTCCGCTGGAGAGACTTGTGTGAGGACACTGTGAGAGACACTGTGACACACCAATTTCCCAACCGGTCA 800
QY 428 TGGCAAAAAATAGAGTTTCTTCCGCTCCAAACAGAGAGAGGCGCGCAAAACAGCTCGCC 487
Db 8002 TGGCAAAAAATAGAGTTTCTTCCGCTCCAAACAGAGAGAGGCGCGCAAAACAGCTCGCC 806
QY 488 TCATCGTATCCACACCTGGAGTTTCGTGTATGCGAGAGATGSGCCCTTTTACGACGTGG 547
Db 8062 TTATCGTATTTCCACACCTGGAGTTTCGTGTGCGAGAGATGSGCTCTCTACGACGTGG 812
QY 548 TTTTCACTCTTCTCTCAGCGCGGTGATGGGCTCTCTCATACGAGTTTCCAAATCTCTCTAAGC 607
Db 812 TTTTCACTCTTCTCTCAGCGCGGTGATGGGCTCTCTCATACGAGTTTCCAAATCTCTCTAAGC 618

Db	8182		AACGGGT	CGAGTTC	CTGTGTGA	CACCTCG	GAATCG	GAAGTGC	CCCTAT	GGCTTCT	CGT	8241							
Qy	668		ATGAC	CACCGCTG	TTTTGAC	CTCAAC	CGGTCA	CTGAGA	ATGACAT	CCGTGTT	GAGGAGTCAA	727							
Db	8242		ATGAC	CACCGCTG	CTTGACT	CAAC	CGGTCA	CTGAGA	TGACAT	CCGTGTT	GAGGAGTCAA	8301							
Qy	728		TTTAC	CAATGTT	TGTGACT	TGCGCC	CCGAA	GCTAG	ACAGG	CCATAG	GTGCTTCA	CAGAGC	787						
Db	8302		TTTAC	CAATGCT	TGTGACT	TGCGCC	CCGAA	GCTAG	CAAA	TTGGCC	ATAAAGT	CGCTCAC	AGC	8361					
Qy	788		GGCTC	TATGTC	GGGGGT	CCCATG	ACTA	ACTCCA	AGGCGA	CACTG	CGGCTAT	GCGCGT	847						
Db	8362		GGCTC	TATGTC	GGGGGT	CCCATG	ACTA	ACTCCA	AGGCGA	CACTG	CGGCTAT	GCGCGT	8421						
Qy	848		GCCG	CGAG	CGGCGT	GCTG	CAGCA	CTAG	CTCG	GGTAA	TACCTC	CACATG	CTACT	TGAAG	907				
Db	8422		GCCG	CGAG	CGGCGT	GCTG	CAGCA	CTAG	CTCG	GGTAA	TACCTC	CACATG	CTACT	TGAAG	8481				
Qy	908		CGCTG	CAGC	CTGT	CGAG	CTGCC	AAAGT	TCC	AGAGT	CTGC	ACATG	CTCGT	CAATG	GAGAGC	967			
Db	8482		CCCTG	CGGCTG	TGCGAG	CTGCA	AAAGT	CTCCG	GACTG	CA	CGATG	CTCGT	GAA	TGGAG	AGC	8541			
Qy	968		ACCTT	GTCG	TTATCT	GTG	GAAG	CGCG	GAAC	CCCAAG	AGAG	CGCG	CAAG	CTTAC	GAGTCT	1027			
Db	8542		ACCTT	GTCG	TTATCT	GTG	GAAG	CGCG	GAAC	CCCAAG	AGAG	CGCG	CAAG	CTTAC	GAGTCT	8601			
Qy	1028		TCACG	AGGCTAT	GACT	AGT	ACTCT	G	CCCCCT	CTGG	GAC	CCG	CCCCA	ACCG	GAATAC	1087			
Db	8602		TCACG	AGGCTAT	GACT	AGT	ACTCT	G	CCCCCT	CTGG	GAC	CCG	CCCCA	ACCG	GAATAC	8661			
Qy	1088		ACTTGG	AGCTG	ATAAC	TCGT	TTCT	CCAA	TGTC	GGT	TCG	GCAC	AGAT	GCAT	CTGG	CA	1147		
Db	8662		ACTTGG	AGCTG	ATAAC	TCGT	TTCT	CCAA	TGTC	GGT	TCG	GCAC	AGAT	GCAT	CTGG	CA	8721		
Qy	1148		AAAGG	GTG	TACT	ACT	CA	CCG	TCG	AC	CCCG	CCCG	CTT	TG	CAG	GGCTG	1207		
Db	8722		AAAGG	GTG	TACT	ACT	CA	CCG	TCG	AC	CCCG	CCCG	CTT	TG	CAG	GGCTG	8781		
Qy	1208		CAGCT	AGGCA	CAC	TCC	AGTCA	CTCT	GGCTAG	GCA	ACATCAT	CATG	TAT	TG	CGCC	CAC	1267		
Db	8782		CAGCT	AGGCA	CAC	TCC	AGTCA	CTCT	GGCTAG	GCA	ACATCAT	CATG	TAT	TG	CGCC	CAC	8841		
Qy	1268		TGTGG	CAAG	ATGAT	CTG	ATG	ACT	CTTCT	TCC	ATCT	CTT	CAT	CTT	CAT	AGCC	CAAC	1327	
Db	8842		TATGG	CAAG	ATGAT	CTT	TGAT	GACT	CA	TTCTT	CTT	CTT	CTT	CTT	CTT	CTT	CTT	8901	
Qy	1328		TTGAA	AAAGCC	TGGAT	TGT	CAG	ATCTA	CGG	GC	TTGTT	ACT	CTC	AT	TG	AGCC	CACT	1387	
Db	8902		TTGAA	AAAGCC	TGGAT	TGT	CAG	ATCTA	CGG	GC	TTGTT	ACT	CTC	AT	TG	AGCC	CACT	8961	
Qy	1388		TACTC	AGATCA	NTG	AAC	GACTCC	AT	GGTCT	TAG	CGCAT	TTT	CAC	TC	CA	TAGT	TA	1447	
Db	8962		TACTC	AGATCA	NTG	AAC	GACTCC	AT	GGTCT	TAG	CGCAT	TTT	CAC	TC	CA	TAGT	TA	9021	
Qy	1448		CAGGT	GAGAT	CAAT	AGG	TGGCTT	CA	TGCTC	AGG	AAAGT	CTG	GGG	TAC	CA	CCCT	CTG	1507	
Db	9022		CAGGT	GAGAT	CAAT	AGG	TGGCTT	CA	TGCTC	AGG	AAAGT	CTG	GGG	TAC	CA	CCCT	CTG	9081	
Qy	1508		TCTGG	AGAC	ATCG	GGCC	CAGA	AGTGT	CGG	GCCT	AAAGT	TACT	GTCC	CA	AGGG	GGG	AGG	1567	
Db	9082		TCTGG	AGAC	ATCG	GGCC	CAGA	AGTGT	CGG	GCCT	AAAGT	TACT	GTCC	CA	AGGG	GGG	AGG	9141	
Qy	1568		CCAT	TTGTG	CAAGT	ACTCT	TTCA	AT	TGG	CA	GTAA	AGG	CC	AAAGT	CTTAA	ACTCA	CTCCAA	1627	
Db	9142		CCAT	TTGTG	CAAGT	ACTCT	TTCA	AT	TGG	CA	GTGG	CA	TAA	AGC	CTTAA	ACTCA	CTCCAA	9201	
Qy	1628		TTCCG	GCTG	GGCTG	GACT	TGT	CGG	CTG	GT	TG	CTG	CTG	CTG	CTG	CTG	CTG	1687	
Db	9202		TTCCG	GCTG	GGCTG	GACT	TGT	CGG	CTG	GT	TG	CTG	CTG	CTG	CTG	CTG	CTG	9261	
Qy	1688		ACATAT	TAC	AGC	CTG	CT	CGT	GGCC	AG	CCCG	CTG	GT	T	CAT	GT	TG	CTCTAC	1747

9262 ACATTATACAGCGTGCTCGTCCGACCCCGCTGTTTCATGTGCTACTCTAC 9332

1748 TCTCCGTAGGGTAGGCATCTATCTACTCCCCAACCG 1784

9322 TCTCCGTAGGGTAGGCATCTACCTGCTCCCCAACCG 9358

RESULT 9

AF207754

LOCUS AF207754 9375 bp RNA linear VRL 27-APR-2000

DEFINITION Hepatitis C virus strain MD13 complete genome.

ACCESSION AF207754

VERSION AF207754.1 GI:7650225

KEYWORDS

SOURCE

ORGANISM

Hepatitis C virus

Hepatitis C virus

Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE

AUTHORS Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Marumo,F. and Sato,C.

TITLE Characteristics of hepatitis C viral genome associated with disease progression

JOURNAL Unpublished

REFERENCE

AUTHORS 2 (bases 1 to 9375)

2 (bases 1 to 9375)

Sato,C.

TITLE Direct Submission

JOURNAL Submitted (23-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan

FEATURES

source

1. 9375

/organism="Hepatitis C virus"

/mol_type="genomic RNA"

/strain="MD13"

/db_xref="taxon:11103"

/notes="isolated from a patient with persistently-normal ALT values (ASC-3)"

330..9362

/codon_start=1

/product="polyprotein"

/protein_id="AAF65944.1"

/db_xref="GI:7650226"

/translation="MSTNPKPQKTKRNNRRPDVKPFGGQIVGGVYLLPRRGPRLVGRTKTSERSQKRGROIIPKARQPRWTQPGYQWPMYVNEGMEGAGWLLSPRGSRPSGPTPRRSNIGKVIDLTGCGADLMGYIPLVCAAGSARALALHGVRLDGNYATGNLPGCSFSIFLIALLSCTIPASAVEVNVSGIVHVTNDCSLTSIVEAADMIMHTPGCVPCVREGNSRCWALTPTLARNSSVPTAIRRHVDLLVGAACFSAFYVGLDGSVFLVSQLTFFSPRRHETTDNCSTIYPGLHTGHRMAMDMNMWTRPTAIVSOLLRTIQAVDMVAGAHGVLGLAYISMAGNWAKVLI VMLLFAGVDGERTVTGCAARATSGVFLSFGSSQNTQLINTNSWHINRTALNCNDSLTQGLAALFYTHKFNASGPERMASCRSIDKFDQGMGPIITYAEPNSDORPCYMHVAPQGIVPASVOCGPVYCFTPSPVGTITDRFGVPTYSWGENITVDLLNTRPPQGNWFGCTMNSTGFTKTCGGPCDGIQVGNNTLTCTCDKFKHPEATYKCGSGPMLTPRCLVNPYRFLMHPYCTNFIPTKVRMYGVGVEHLEAKCANTRGERCDLDDRSSELSPLLLSTTEQYIPLCSFTLLPALSTGLILHNNIYVDVQYLYIGSASVFAIKWEYVLLFLLLADARVCAALNMWLLIAQASAALENVLVNAASVAGTRGILSLFVFCFAAWYIKGRILVPCAAYALGVMPLLMLLALLPPRAYANDRWSGCGAVFGLALLTLSPYKVFARIIWNQYILITRAEHALHWVPPNLNVRGDRDAILLTCAVHPELIFDITLLAILGLPMVFORGITRVPYFVARQGLIRACMLRKVVYGVQVAFKMAALTGTYYVDHLTPURDWAHTGLRDLAVAVEAYSDQTMETKIIITWAGDTAAACDIIISGLPVSARRGKEILLGPADSLGEGWLLAPITAYSVQTRGMJGCIITSLTRKDNQVGEVQVSTATOSFLATCVNAGCMTVYHGAGSKTLAGPKGPIITOMYNTVDLQWQAPPGARSILTPCTCGSSDLYLVRHADVI PVRRRGDERGSLSPRPVSVYKSGSGPLLCPSGVGVI FRAAVTRGVAADVFPVESMETTKRSPVFTNSPSPVAPVQFTGVGHLPHTGSGKSTKVPAAAYAOGYKVLVNLNPSVAATLFGFAYNSKAGHGVDPNIRTGATITTGAPITYSTYKFLADCGSCGQVDIIICPECHSTGDTTILGTLVDOAETAGARLVLTATAPPQSVTPPHNNIEAALSTGEIIPPYKAIPIETIKGGRHLIFCHSKKCDLAAKLSGLGNNAVYRGLDVSIVPTSGDVVVVA TALMNTGYTGPDSVDNCTVQTVDFSLDPTIETITTVFDVAVSRQRGRTGRGRGIYRFTVGPDSVCLCECDAGCAWELTPAETSVRLRAYLTNTPGLPVQDHLFEWSEFTGLTHIDAHFISQTKAGDNFPYLVAOYVCARQAQAPPSQDMWCKLTRLRPLTGHGPTLLYLRGA VONEVTLTHPTIKTYIMACMSADI EVMVSTWYAGCVTA

ALAAAYCLTTGSSVVI VGR IILISQRPAV I PDREVLYR BFDMEECASHU PYI TEQGMQJAE
 OFKQKAGLLTQAT PQTAAAAAPAVESKRNKALBETVWAKHMWNF I SGI QYLAGLSTFPGN
 PA IASIAWALTSQT KLTAAAAATVIR LCGWAAQALVPAASAFAGAGAGAGAGS I
 CLGVYVD IILAGYAGAGAGALVAGWNSGTSPHYEESALVNTLPA I LSPCALVWGWCAA
 LLRRHVGSGAGVQWNNL IAFASRGNHVSPTHYEPESALVNTLPSLITLQLLK
 RLHOW INEDCSZSGSGWURDWNMI LCTVADFKALWSKULPAGVVPFSCQRYK
 GWNFGDQIMOTTCPCQA ITHGVYGHGVMITGVCTSNKTHGTFPI INAYITGCVTVLSP
 APNTSRALVRAAEYEVYTVRGDFHVTMTGTVDNVQPCQVPABEFTVTPDGVTRH
 YAPACKPULREYTVFVGLMJO I VGSOLPCEBPEDVATVLSMLDTPSTAEATKARL
 ARGSPSLAESASOLASPSILATCTTHHDSPADL I EANLWQEMGNI TVRSEEN
 KVIILBSFDP I RAEDEREVSVPAE I LRKSRFEPAL I PAWPNYNPVLESWKOPDY
 VPPVHGQSDPTTAPPI PPRKRTIUVLITESVSAALAEALATVTPGSSSESSAADSGT
 VAPVDAQSDGDKESDYESVSMPPLEGEQDPDLSGMSNTVSEASESDVCCSSMS
 YVTWGTALITPCAAEESKLPI NPLNSLRLHHNNVATTSRAGLRKQVTFDRGLVNI
 DHYEVD I KEMKATASVKALMSLVEZACKLT I PHSKAGSKFGYGAKRDNVLSQAVHI
 RSVNKLLEDTEPT I DIT I WAKNEVECPQEPKSKRPAR I TVFPDLGVYRCEKMAIYD
 VYSTLPQAVGSSYGFQYS PQWSEVETLWMTSKSKCPWGESCTPSTVTFESDIRV
 BESTYOCDDLAPEAKI RSLTEL I YIGPLTNSGQNGCVRR CWSAGLTVTCSGNTL
 TCYUKATCAAAKALQDCTMLVNGDDVL I CESSAGTQDPAALAKRAFTAMTRYSAPPG
 DLQPEYDLEL I TSCSNVSNAGDESKRKYVITRDPITVPLAARAEATRH PWNVSLW
 GNIIMVYPTWIRNVLWTFES I LLAQEOLAKDQCYAGKSY I EPIDLPQ I IERLH
 GLSIAFTSPYSGE I NVASCL I LAGQVPLVRWHRASVRAKILSSQGRRAATCGKY I
 FNWAVRVTCLLTTPNPAASRLDLSGWTFVAGVNGCD I VHSLSRARPWFMLCLLLSSVGV
 FNWAVRVTCLLTTPNPAASRLDLSGWTFVAGVNGCD I VHSLSRARPWFMLCLLLSSVGV

CONCLUSIONS

Query Match	89.7%	Score 1617.8	DB 14	Length 9375	
Best Local Similarity	94.5%	Pred. No. 0			
Matches 1676	Conservative	0	Mismatches 97	Indels 0	Gaps 0
Qy	13	TCAATGTCCTATACATGACACAGGCGCCCTGATCACACGTCGGCTGGGAGGAAGCAAG	72		
Db	7587	TCAATGTCTACATGACACAGGTGCCCCGTGATCAACCATCGCTGGGAGGAGCAAA	7646		
Qy	73	CTGCCCATCAACGGCGCTGAGCAACTCTTGTCTGGTCACCATACCTTGGTCTATTCCACA	132		
Db	7647	CTGCCCATCAACCCGTTGAGCAACTCTTGTCTGCTCACCACACATGTGTGTATGCCACA	7706		
Qy	133	ACATCCCCAGTGCAAAGCCTCGGCGAGAAGAGGTCACTTTGACAGACTCAAGTCCTG	192		
Db	7707	ACATCCCGAGTGCAGGCCCTGCGGCGAGAAGAGGTCACTTTGACAGACTCAAGTCCTG	7766		
Qy	193	GACGATCATTTACCGGGACGTGCTCAAGCAGATGAAGCGCAAGGCGTCCACAGTGAAGCT	252		
Db	7767	GACGACCACTACCGGGACGTGCTCAAGGAGATGAAGGCGAAGGCGTCCACGTTTAAGCT	7826		
Qy	253	AAACTGCTATCTGTAGAAGAAGCATGCAAGCTGACGCCCCCGCATTTGGGCCAAATCCAAA	312		
Db	7827	AAACTTCTATCTGTAGAAGAGCCTGTAACTGACGCCCCCACAATTCGGCCAAATCCAAA	7886		
Qy	313	TTTGGCTATGGGGCAAAGSACGTCGGGAGCCTATCCAGAGGGCGCGTTAACACCATCCGC	372		
Db	7887	TTTGGCTACGGGGCAAAGSACGTCGGGAACCTATCCGGCAGGCGGTTAACACCATCCGC	7946		
Qy	373	TCCGTGTGAAGGACTTGTGTGAGGACACTGACACCAATTCAGACCACCATCATGGCA	432		
Db	7947	TCCGTGTGAAGGACTTGTGTGAAGACACTGAGACCAATTTGATACCAACCATCATGGCA	8006		
Qy	433	AAAAATGAGGTTTTCTGCGTCCAAACAGAGAAGGAGGCGCAAAACCAAGCTCGCCTCATC	492		
Db	8007	AAAAATGAGGTTTTCTGCGTCCNAACAGAGAAGGAGGCGCAAGCCAGCTCGCCTTATC	8066		
Qy	493	GTATTTCCACAGACTGGGAGTTTGTGTATCGCAGAGATGGCCCTTTTACAGCTGTGTTTCC	552		
Db	8067	GTATTTCCACAGACTGGGAGTTTGTGTATCGCAGAGATGGCTCTTTTACGACGTGTGTTCC	8126		
Qy	553	ACTCTTCTCAGGCCGTGTGGGCTCCTCATACGGATTCCAATACTCTCTTAAGCAGCGG	612		

QY	1753	GTAGGGGTAGGCATCTACTTACTTCCCAACCGG	1785	
Db	9327	GTGGGGTAGGCATTTTACCTGCTCCCAACCGG	9359	
RESULT 10				
LOCUS	AF207767	9379 bp	RNA linear	VRL 27-APR-2000
DEFINITION	Hepatitis C virus strain MD26 complete genome.			
ACCESSION	AF207767			
VERSION	AF207767.1	GI:7650251		
KEYWORDS				
SOURCE	Hepatitis C virus			
ORGANISM	Hepatitis C virus			
REFERENCE	Viruses: sRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			
AUTHORS	1 (bases 1 to 9379) Nagayama, K., Kuroesaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and Sato, C.			
TITLE	Characteristics of hepatitis C viral genome associated with disease progression			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 9379) Nagayama, K., Kuroesaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and Sato, C.			
AUTHORS	Direct Submission			
TITLE	Submitted (23-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan			
JOURNAL	Location/Qualifiers			
REFERENCE	1..9379			
AUTHORS	/organism="Hepatitis C virus"			
TITLE	/mol_type="genomic RNA"			
JOURNAL	/strain="MD26"			
FEATURES	/db_xref="taxon:11103"			
source	/note="isolated from a patient with liver cirrhosis and hepatocellular carcinoma (HCC-6)"			
	330..9362			
	/codon_start=1			
	/product="polyprotein"			
CDS	/protein_id="AAF65957.1"			
	/db_xref="GI:7650252"			
	/translation="MSTNPKPQRKTKNWRPDPVKFPGGGQIVGGVLLPRRGRPL SRPTKRSRSGRQRIPKAPQEGRTWAQPGVPWPLVNGEWMGWLSPRG GVRSGPTDPRRRNLGKVIDTLTCGFADLMGYIPLVGAPLGGARALAHGVRLD SWNATGNLPCGSFIFLLALSLCTLPASAEVRNVSGIYHVNDCNSSIYEAAD IIMHAPGTCPCVRSGKSRCHWALTPLIARNISVPTTIRRHVDLLVGAAPFCSAMY VGLDCGVFLVQLTFLSPRRQHTQDNCISYPGHVSGRHWADMMNNSPTAALVY SQLRIIPQAIMDVMVGAHWGLAYYSMVGNAKVLIVMLIPAGVDGDTRTVGAS AHNRTTTSFTSPGSKQIOLVNTNGSMHNRNALNCDSLOTFLAALFYTHRNAS GCPERMAACRPIDNPDQGWGITVAEPNSPDORPCYHVALRPGCTTPASGCVGPYC FTPSVVGTTDRFGAPTYNGENETDILNNTPLPQGNWFGCTWMSNFTGTXCTGG PPCNLGAGNNTLICPTDPRKPEATYTKGSGPWLTPRCLVDYPYRLMHPYPTVF TIFKRVYVGEHRSUACNWRGECGLDRDSELSFLLSTTEWOILPGSFITL PALSTGIHNRHVDVQYILGCSAVSPFAIKWEYVLLFLLADARVACALNWMLL VAQEAALNVLNMAASVAGHGILSFVFCFAWYIKGRVLPGBAVALGVWFLLL LLALPRAVALREMAASCGGVFLGALLLILSPYKVFARL.IWLOIFITRAEAB LQWIPPLNVGRGDAILLVCAVHPFELIPDITKLALLIFGLMVLQASITRVYFVR AQGLIRACMLRVAGHYVOMALVKLAALTGYTYVDHLAPLQHWHAHLRLAVAVE PVFSDMETKIIITWGDATACQDIIILGLPVSARRGREIILGPADEFGQWELLAPIT AYSQOTRGLGLCIIITSLTRDKNQVEGQVWSTQSFATCVNGVCMVTFHGAGSK TLAQPKGPIQMTYTNVDOLVQWQAPPGARSMPTCTCGSSDYLIVTRHADVIPVRRG NBSGSLSPRPISYLKSGSGGPLCPLGHHVGI FRAACVTRGVAKAVDFVPVSEMT DRSPVFTDNGSPVAFPTFOVAHLHAPTSGSKTKVPAAYAAQGYKVLVLPNSAATL GFGYMSKAYGTDPNRTGIRTTTGAITYSTYKFLADGCGGAYDIIICDECHS TDSITLIGITGVLDQAEATAGARLVILATTPGSVTVPHNIEEVALSNTGEIIPYK AIPLEVIKGRHLIFCHSKKCDLAKLSALGINAVAYRGLVDSVITPGSDVVVA TDALMTCTGDFDSVIDONTCTVOTVDFSLDPTFTLETITTVQDARSQRGRTRG RIGYIRVTFEGRPSGMSVJCECTDAGCAWELTPSETTVRLRAYLNTPGIPVCO DHLEFWSVETGLTHDHFSLQTKOAGNFPYLVAYQATCARAQAPPPQWDMWKC LILRLPTLPTGLVRLGAVONEVLTHTPIKTFTIMCAADLEVVTSTVYLVGVLA ALAAVCLTGTGSVVIGRIILSGKPAIPDREVLREFDEMEBESALSTVIEQQLQAE QFKOKALGLLQATKQAEAAAPAVESKREALFAFWAKHWNFIISGILYLAGLSLGN			

Db	8242	ATGACACCCGCTGTTTGGACTCAACGGTCAACGAGTGACATCCGTTACTGAGAGTCAA	8301
QY	728	TTTACCAATGTTGTGACTTTGGCCCCGAAGCTAGACAGCGCCATTAAGTTCGCTCACAGAGC	787
Db	8302	TTTACCAATGTTGTGACTTTGGCCCCGAAGCTAGACAGCGCCATTAAGTTCGCTCACAGAGC	8361
QY	788	GGCTCTATGTCCGGGGTCCCATGACTAACTCAAAGGCGAGAACTGCGGCTATCGCCGGT	847
Db	8362	GGCTCTATGTCCGGGGTCCCATGACTAACTCAAAGGCGAGAACTGCGGCTATCGCCGGT	8421
QY	848	GCCGCGCAGAGCGGCTGTGAGGACTAGTGGTGTAACTACCTCACATGCTTACTTGAAGG	907
Db	8422	GCCGCGCAGAGCGGCTGTGAGGACTAGTGGTGTAACTACCTCACATGCTTACTTGAAGG	8481
QY	908	CCGCTGCGAGCCTGTGAGGCTGCCAAGCTCCAGGACTGCACAGATGCTGTGAATGGAGACG	967
Db	8482	COACTGCGGCTGTGAGGCTGCCAAGCTCCAGGACTGCACAGATGCTGTGAATGGAGACG	8541
QY	968	ACCTTGTGCTTATCTGTGAAAGCGCGGAACCCAAAGAGGAGCGCGCAAGCCTACGAGTCT	1027
Db	8542	ACCTTGTGCTTATCTGTGAAAGCGCGGAACCCAAAGAGGAGTGCAGGAGCCTACGAGTCT	8601
QY	1028	TCACGAGGCTATGACTAGGTAATCTGCCCCCTTGGGGACCCGCCCCCAACCGGAATACG	1087
Db	8602	TCACGAGGCTATGACTAGGTAATCTGCCCCCTTGGGGACCCGCCCCCAACCGGAATACG	8661
QY	1088	ACTTGAGGCTGATACATCATGTTCTCCATGTTGCTGCTGCCACACCATGCATCTGSCA	1147
Db	8662	ACTTGAGGCTGATACATCATGTTCTCCATGTTGCTGCTGCCACACCATGCATCTGSCA	8721
QY	1148	RAAGGCTGACTACTCTACCCGCTGACCCACCGTCCCTTGGCGGGCTGCGTGGGAGA	1207
Db	8722	RAAGGCTGACTACTCTACCCGCTGACCCACCGTCCCTTGGCGGGCTGCGTGGGAGA	8781
QY	1208	CAGCTAGGCACACTCCAGTCAACTCTGCTGCTAGGCAACATCATGATGTATGCGGCCACTT	1267
Db	8782	CAGCTAGGCACACTCCAGTCAACTCTGCTGCTAGGCAACATCATGATGTATGCGGCCACTT	8841
QY	1268	TGTGGGCAAGATGATCTTGATGACTCATCTCTTCTCCATCTCTTAGCCCGAGGAGCAAC	1327
Db	8842	TATGGGCAAGATGATCTTGATGACTCATCTCTTCTCCATCTCTTAGCCCGAGGAGCAAC	8901
QY	1328	TTGAAAAGCCCTGGATTCAGATCTACGGGCTTGTACTTCCATTCAGGCACTTGACC	1387
Db	8902	TTGAAAAGCCCTGGATTCAGATCTACGGGCTTGTACTTCCATTCAGGCACTTGACC	8961
QY	1388	TACCTCAGATCATTTGAACGACTCCATGCTTCTAGGCACTTTTCACTCCAATGTTACTCTC	1447
Db	8962	TACCTCAGATCATTTGAACGACTCCATGCTTCTAGGCACTTTTCACTCCAATGTTACTCTC	9021
QY	1448	CAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACACCCCTTGGGAG	1507
Db	9022	CAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACACCCCTTGGGAG	9081
QY	1508	TTGAGAGACATCGGGCCAGAGTGTCCGGCTTAAGTTACTGTCTCCAGGGGGGAGGGCGG	1567
Db	9082	TCTGGAGACATCGGGCCAGAGTGTCCGGCTTAAGTTACTGTCTCCAGGGGGGAGGGCGG	9141
QY	1568	CCATTTGTGCAAGTACCTCTTCAACTGGGCACTAAGGACCAAGCTTAAACTCACTCCAA	1627
Db	9142	CCATTTGTGCAAGTACCTCTTCAACTGGGCACTAAGGACCAAGCTTAAACTCACTCCAA	9201
QY	1628	TTCCGGCTGCTCCCGGCTGGCACTTGTCCGGCTGGTTCGTTGCTGCTACAGCGGGGAG	1687
Db	9202	TCCGGCTGCTCCCGGCTGGCACTTGTCCGGCTGGTTCGTTGCTGCTACAGCGGGGAG	9261
QY	1688	ACATATATCAGAGCTGCTGCTGCGCGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCT	1747
Db	9262	ACATATATCAGAGCTGCTGCTGCGCGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCT	9321
QY	1748	TCTCCGTAGGGTAGGATCTATCTACTCCCAACCG	1784

Db

9322

TTTCCGTAGGGTAGGATCTATCTACTCCCAACCG

9358

RESULT 11

HPCRNA

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

1

(bases 1 to 9456)

Wang,Y., Okamoto,H., Tsuda,F., Nagayama,R., Tao,Q.M. and Mishiro,S.

Prevalence, genotypes, and an isolate (HC-C2) of hepatitis C virus

in Chinese patients with liver disease

J. Med. Virol. 40 (3), 254-260 (1993)

93359897

8394876

These data kindly submitted in computer readable form by: Hiroaki Okamoto

Immunology Division

Jichi Medical School

Kawachi-gun

Tochigi 329-04

Japan

Phone:

0285-44-2111 x3334

0285-44-1557.

Location/Qualifiers

1. .9456

/organism="Hepatitis C virus"

/mol_type="genomic RNA"

/isolate="HC-C2"

/db_xref="taxon:11103"

1. .341

342. .9374

/note="C, E, NS1/E2, NS3, NS4 and NS5"

/codon_start=1

/product="polyprotein precursor"

/protein_id="BA01728.1"

/db_xref="GI:471117"

/translation="MSTNPKPQRKTNTNRRPDVDFKPGGQIVCGVYLIPRRGPRP

GVTRKTSRSQPRRRQPIPKARPEGRWAQPGYPWPPLYGNEGMGWAGWLLISPR

SRPWSGNDPRRSRLGKVIDTLTCGADLMGYIPLVAPLGAARALAGVRALE

GVNATCNLPGCSFIFLLALLSLTIPASAEVVRNASGVHVNTDNCNSIIVVEAA

MINHPGCPVCRENNSSRCWVALPTLAARNASVPTTIRHVLLVGAALCSAMV

VGDLGGSFVLVSQLTFTSPRRHETVQCNCIYPGHITGHRMAMWMMNSPTTALV

SQLRIPQAVMDVMVAGAHWGLAGLAYYSWGNWAKVLIIVLLIPAGVDGNTYVTGGAP

ARGASGITSLFSDRPSQKIQLVNTNGSHINRTALNCNDSFNTGFLAALFYAHRFNS

GCPRMASCISDKFDQWGPITYQDSDQRPYCWHPYPRPGIIVPASEVCGPVV

PTSPVYVGTDRLGVTVMNENETDVLNNTNRPQGNWFGCTWMTTGTCTCG

PPCNIGGAGNLTCTPDCFRKHPEATYKCGSGPALTPLCLVDYVYRLWHYPCTVNI

TIFKRYMVGVEHRLDAACNWTGRCALEDNRSELSPLDJSTTEWQLILPCSFTHI

PALSTGLIHLHNRIVDQYLYGIGSAVSFAIKWEYVLLILLADARVACALMMWLL

IAQAEALENLAVLNAAASVAGVHGILSPLVFFCAAWIKGRIVPAAYAFYVWPLLI

LIILALPPRAYAMAREMAASCGAVFGLALLTILSPYKVLARLIWLOVFTVRAEA

LOVMIPLNVRGRDAIILLTCAHPELIIFDIITKLALAIEGPTWVLOAGITKMPVFV

AOGLIRACMLVRKVGAGHYQMAEMKAAULTGTIVYDHLIFLRDWAHTGLRLDAVAVI

PVWFSDMETKIITWGADTAACDIIILGLPVSARRGKEIILGPADSLBGQGRLLIPI

TDSITLIGITVLDQAEATAGARLVVLAATATPAGSVTPVSHYSLNEEVALSTGTEIIPYGH

GFGYMSKAGHIDINTRTGVRTITTPGITTYSTYTKFLADGCGCGGAYDIIICDECH

AIPIETIKGGRHLIFCHSKKCKDELAALGNVAYVYHNSRISVEALSTGTEIIPYGH

TDALMTGTFGDSVIDCNTCTVOTVDFSLDPTFTTETTVDPQAVSRQRGRITGR

ARGIYRFTVTPGERSGMFDSSVLCEYDAGCAWVELTPTAETSVRLRAYLNTPLGIPC

DHLEPWSVFTGLTHIDAHFVSQTKQAGDNFPLVAYQATYCARAQAPPPSDQNMWK

LIRLKLPLHGTPTLLYRIGAVONEVTLTHPIIKYIMACMSADLEVETWVTLVGVGL

ALAAAYCLTTGSVIIVGRIVLSGRPAVIFDREVLYQEFDEMBEBCGSHLPIYISQGMOLA

FEATURES

source

5'UTR

CDS

OPKOKALGLOIATKQAEAAA PWESKWRALFTFWAKMWNFI SGVOYLGLSLPGN
PALASIMAFASVTSPLTTQSTLLFNILGCVAAQLAPPSAASAPV GAGIAGAAGVSI
GLGKVLVDLAGVAGAGALVAFKIMSGTSPAEEDLVNLLPAILSLVGVVVCAA
IQRRHVGPGEVQWNRNLIAPASGNHVSPTHYFESDAAARVTOILSLITQLLK
RLHWINEDCSTPCSGNWLRIWDMICSVLDTFKWQSLKLLPRLPVFPFSCORGYK
GWRGDIMQTTCPQCAQITGHVKNIGSMRI VGPRTCSNTHGTFPFINAYTTGCPCTP
APNYSALWRAAEVEVETRVGDHYITGMNTDNI KCPQVPAEPETVEDVGPLHR
YAPACKPVLREVDVQVLGNQYVPGSQLPCPEPDVAVLTSMLTDFSHIAPAAKRL
AKGSPSLASSASQLSAPLTKATCTTHHSDPADLI BANLLWRMGENTIRVESEN
KVYILDSFPLRAEDEREVSAAHILKTRTFPPAMPVYMARPDYNPBLLSWKDPDY
VPPVHGCPPLPTKAPPPIPPRRKRTVLVTSTVSSALAEIATKTFSGSGSAVDST
ATAPDOTSNDGSDRESDAESYSPMPLEGEQD PDLSDGSWSTVSEASGDVCCSMS
YTMTCALITPCAARESKULPINLSNLLRHNMVYATTSRSASLRKKYTFDRLOVLD
DHYRVLKMKAKASTYKALJUSIEACKLTPHSKSKFPGYKANDVRNLSKAVNHI
RSVWDLDETPIDITIMAKNEVFPQVPERGKRPALIVFPDLGVYVCRKMALYD
VYSTLPQVMGSSYSPQSPQRFVFLVNTWKSKKCPMGFSYDTRCPSTVENDIRY
BESIYQCCDLAPAEQAOKSLVLYIGGLPNSKQNGCYRRCSRASGLVJTSSCQNTL
TCYLKASACRAAKLQDQTMVLVNGDVLVVICESACTORDAASLRVFTAMTRYSPAG
DPQPEYDELITSCSSNVSAHDASGRVYLYLTRDPTPLABAANETARHTPPVNSWL
GNITAPTLWARMILMTHFFSILLIAQOLEKALBCQIYGACTSIEPDLQIITRLH
GUSATLSHSYSPGBINRVASCLRLGVPRLVYMRHRARSVRKLLSQSGRAATCKYL
FNWAVRTKLTPIPAAASRLDLSGWFVAGYGGGDIYHSLSRAPRPWFMLCULLLSVG
GYLLPNR"
3'UTR 9375.. 9456
/note="with a T-stretch of 45 bp."

ORIGIN

Query Match	89.5%;	Score 1613.8;	DB 14;	Length 9456;	
Best Local Similarity	94.3%;	Pred. No. 0;	Mismatches 102;	Indels 0;	Gaps 0;
Matches 1675;	Conservative 0;				
QY	8	GCATGTCAATGTCCTATACATGACAGCGGCCCTGATCACACCGTGGCTGCGGAGGAAA	67		
DB	7594	GCTGCTCAATGTCCTACATGACAGAGTGCCCTTAATCACACCATGCGCTGCGGAGGAAA	7653		
QY	68	GCAAGTCGCCATCAACCGCTGAGCAACTCCTTGTGCGTCAACATAACTGGTCTTATT	127		
DB	7654	GCAAGTCGCTATCAACCTCTGAGCAACTCCTTGTCTGCGTCAACCAACATGGTGTATG	7713		
QY	128	CCACACATCCCGAGTGAAGCCTCGGCGAGAGAGGTCACTTTGACAGACTGCAAG	187		
DB	7714	CCAACAATCTCGCAGCGCAAGCTCGGCGAGAGAGGTCACTTTTGACAGATTGCGAG	7773		
QY	188	TCTCGACGATCATTAACCGGAGCTGCTCAAGAGAGATGAAGCGAGGCGTCCACAGTGA	247		
DB	7774	TCTCGAGCACCATTACCGGAGCTGCTCAAGAGAGATGAAGCGTCCACAGTTA	7833		
QY	248	AGGCTAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCCCGCATTCGCGCAAAAT	307		
DB	7834	AGGCTAACTTCTGTCTATAGAAGAGCTGCAAGCTGACGCCCGCCACATTCGCGCAAAAT	7893		
QY	308	CCAAATTTGGCTATGGGCAAGAGAGCTCGGAGCCTATTCAGAGGCGCGGTTAACCA	367		
DB	7894	CCAAATTTGGCTATGGGCGAAGAGAGCTCGGAAACCTATCCAGCAAGGCGGTTAACCA	7953		
QY	368	TCCGCTCCGCTGGAAGGACTTCTCGAGGACACTGACACACCAATTCAGACCACTCA	427		
DB	7954	TCCGCTCCGCTGGAAGGACTTCTGGAAGACACTGAAACCAATTCAGACTACCATCA	8013		
QY	428	TGGCAAAATGAGGTTTCTCGCTCCAAACAGAGAAAGAGGCGCGCAACACAGCTGCGC	487		
DB	8014	TGGCAAAAGATGAGGTTTCTGCGTCCAAACAGAGAAAGGAGGCGCGCAAGCCAGCTGCGC	8073		
QY	488	TCATCGTATTTCCAGACCTGGGAGTTGCTGTATGCGAGAGATGGCCCTTTACGAGCTGG	547		
DB	8074	TTATCGTATTTCCAGACCTGGGGGTTCTGTGTATGCGAGAGATGGGCTCTTTATGAGCTGG	8133		
QY	548	TTTCACACTCTTCTCGAGCGGCTGATGGGCTCTCATACGATTCCAATACTCTCTCTAAGC	607		
DB	8134	TCTCCACCTTCTCTAGGCGGTGATGGGCTCTCTATACGATTTCCAGTACTCTCTCTGGGC	8193		
QY	608	AGCGGTCGAGTTCTCTGGTGAATACCTTGGAAAGCAAGAAATGCCCTTATGGGCTTCTCAT	667		

DB	8194	AGCGGGTCGAGTCTCTGCTGTAATACCTGGAATCAAGAAATGCCCTATGGCTTCTCAT	8253
QY	668	ATGACACCGCTGTTTGGACTCAACGGGTCACTGAGAATGACATCCGTTTGAGGAGTCAA	727
DB	8254	ATGACACCGCTGTTTGGACTCAACGGGTCAACGGGTCAACGGGTCAACGGGTCAAC	8313
QY	728	TTTACCAATGTTGACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	787
DB	8314	TTTACCAATGTTGACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	8373
QY	788	GGCTCTATGCTCGGGGTCCTCACTAATCTCAAGGGGCAAGCTGCGGCTATCCCGCT	847
DB	8374	GGCTCTATGCTCGGGGTCCTCACTAATCTCAAGGGGCAAGCTGCGGCTATCCCGCT	8433
QY	848	CGCGCGAGCGGGGTCGTGACGACTAGCTCGGCTAATACCTCACATGCTACTTTGAAGG	907
DB	8434	CGCGCGAGCGGGGTCGTGACGACTAGCTCGGCTAATACCTCACATGCTACTTTGAAGG	8493
QY	908	CGCTGACGCTGTCGAGCTCCAGCTCCAGGACTGACGATGCTGCTGTAATGAGAGC	967
DB	8494	CCTCTGACGCTGTCGAGCTCCAGGACTCCAGGACTGACGATGCTGCTGTAATGAGAGC	8553
QY	968	ACCTTGTGTTATCTGTGAAGCGGGGAAACCCAGAGGATGCGGCGAGCTACGAGCT	1027
DB	8554	ACCTTGTGTTATCTGTGAAGCGGGGAAACCCAGAGGATGCGGCGAGCTACGAGCT	8613
QY	1028	TCAGGAGCTATGACTAGGTACTCTGCCCTCCCTGCGGACCCGCCCAACCGGAATACG	1087
DB	8614	TCAGGAGCTATGACTAGGTACTCTGCCCTCCCTGCGGACCCGCCCAACCGGAATACG	8673
QY	1088	ACTTGGAGCTGATAACATCGTGTCTCTCAATGTGCTGCTGCTGCTGCTGCTGCTGCTG	1147
DB	8674	ACTTGGAGCTGATAACATCGTGTCTCTCAATGTGCTGCTGCTGCTGCTGCTGCTGCTG	8733
QY	1148	AAAGGGTGTACTACTCCTCACCCTGACCCGCTCCCTCCCTGCGGCTGCTGCTGCTGCTG	1207
DB	8734	AAAGGGTGTACTACTCCTCACCCTGACCCGCTCCCTCCCTGCGGCTGCTGCTGCTGCTG	8793
QY	1208	CAGTACGACACTCCAGTCAACTCCTGGCTAGGCAACATCATGATGATGCGGCTT	1267
DB	8794	CAGTACGACACTCCAGTCAACTCCTGGCTAGGCAACATCATGATGATGCGGCTT	8853
QY	1268	TGTGGGCAAGATGATCTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1327
DB	8854	TATGGGCAAGATGATCTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	8913
QY	1328	TTGAAAAGCCCTGGATTTGTCAGATCTTACGGGGCTTGTACTCCATTGAGCCACTTGACC	1387
DB	8914	TTGAAAAGCCCTTAGAGTGTAGAGTGTAGAGGCTTGTACTCCATTGAAACCGCTTGACC	8973
QY	1388	TACCTCAGATCATTTGAAGGACTCCATGCTTCTAGGCAATTTTCTTCTTCTTCTTCTTCTT	1447
DB	8974	TACCTCAGATCATTTGAAGGACTCCATGCTTCTAGGCAATTTTCTTCTTCTTCTTCTTCTT	9033
QY	1448	CAGGTGAGATCAATAGGGTGGCTTCTATGCTCAGGAGCTTGGGCTACCCCTTGGCAG	1507
DB	9034	CAGGTGAGATCAATAGGGTGGCTTCTATGCTCAGGAGCTTGGGCTACCCCTTGGCAG	9093
QY	1508	TCTGAGACATTCGGGCGCAGAGTGTCCGCGCTAAGTTACTGTCTCCAGGGGGGAGGGCGG	1567
DB	9094	TCTGAGACATTCGGGCGCAGAGTGTCCGCGCTAAGTTACTGTCTCCAGGGGGGAGGGCGG	9153
QY	1568	CCATTTGCGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAACTACTCTCAA	1627
DB	9154	CCATTTGCGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAACTACTCTCAA	9213
QY	1628	TTCCGGCTGCTCCGGCTGAGCTTGTCCGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1687
DB	9214	TTCCGGCTGCTCCGGCTGAGCTTGTCCGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTG	9273
QY	1688	ACATATATCAGAGCTGTCTGCTGCCGACCCCGCTGGTTCATGTTGCTGCTTCTCTTAC	1747
DB	9274	ACATATATCAGAGCTGTCTGCTGCCGACCCCGCTGGTTCATGTTGCTGCTTCTCTTAC	9333

Db 8182 AGCGGTCGAGTTCCTAGTGAATGCTTGAATCAAGAAATGCCCTATGGCTTTTCAT 8241
Qy 668 ATGACACCGCTGTTTGTGCTCAACGGTCACTGAGAAATGACATCCGTGTGAGGAGTCAA 727
Db 8242 ATGACACCGCTGTTTGTGCTCAACGGTCACTGAGAGTGACATCCGTGTGAGGAGTCAA 8301
Qy 728 TTTACCAATGTTGCTGAGTGGCCCGGAGCTAGACAGGCCAATAGTCCGTACAGAGC 787
Db 8302 TCTACAGAGTGTGCTGACTTGGCCCGGAGCCAGACAGGCCAATAGTCCGTACAGAGC 8361
Qy 788 GGCCTCTATGTCGGGGTCCCATGACTTAACCTCAAAAGGGCAGAACTGCGGCTATCCCGGT 847
Db 8362 GGCCTCTATGTCGGGGTCCCATGACTTAACCTCAAAAGGGCAGAACTGCGGCTATCCCGGT 8421
Qy 848 GCGCGCGAGCGCGTGTGAGCTAGCTGCGGTAAATACCTCACATGCTACTTTGAAGG 907
Db 8422 GCGCGCGAGCGCGTGTGAGCTAGCTGCGGTAAATACCTCACATGCTACTTTGAAGG 8481
Qy 908 CGCTGCGAGCTGTGAGCTGCGAGCTCCAGGCTGCGAGCTGCGAGCTGCGAGCTGCGAGC 967
Db 8482 CCTCTGCGAGCTGTGAGCTGCGAGCTCCAGGCTGCGAGCTGCGAGCTGCGAGC 8541
Qy 968 ACCTTCTGCTTATCTGTGAAGCGCGGAGCCCAAGAGAGCGCGGAGCGGAGCTGCGAGCT 1027
Db 8542 ACCTTCTGCTTATCTGTGAAGCGCGGAGCCCAAGAGAGCGCGGAGCGGAGCTGCGAGCT 8601
Qy 1028 TCACGAGGCTATGACTAGTACTCTGCCCCCTTGGGAGCGCGGAGCGGAGCGGAGCTGCGAGCT 1087
Db 8602 TCACGAGGCTATGACTAGTACTCTGCCCCCTTGGGAGCGCGGAGCGGAGCGGAGCTGCGAGCT 8661
Qy 1088 ACTTGGAGCTGATAACATCGTGTCTTCCATGCTGCGTGCAGCAGCATGCTATGCGCA 1147
Db 8662 ACTTGGAGCTGATAACATCGTGTCTTCCATGCTGCGTGCAGCAGCATGCTATGCGCA 8721
Qy 1148 AAGGGTGTACTACTTCACTCCCTGACCCCGTCCCGTCCCGTCCCGTCCCGTCCCGTCCCGT 1207
Db 8722 AAGGGTGTACTACTTCACTCCCTGACCCCGTCCCGTCCCGTCCCGTCCCGTCCCGTCCCGT 8781
Qy 1208 CAGCTAGGACACTCCAGTCACTTCTGCTAGGCAACATCATCATGATGATGCGGCCACTT 1267
Db 8782 CAGCTAGGACACTCCAGTCACTTCTGCTAGGCAACATCATCATGATGATGCGGCCACTT 8841
Qy 1268 TGTGGCAAGGATGATCTGTGATGACTCACTTCTTCCATCTTCTAGCCAGGAGCAAC 1327
Db 8842 TGTGGCAAGGATGATCTGTGATGACTCACTTCTTCCATCTTCTAGCCAGGAGCAAC 8901
Qy 1328 TTGAAAAAGCCCTGGATTGTGAGTCTACGGGCTTGTACTCCATTGAGCCACTTGACC 1387
Db 8902 TTGAAAAAGCCCTGGATTGTGAGTCTACGGGCTTGTACTCCATTGAGCCACTTGACC 8961
Qy 1388 TACCTCAGATCATTTGAACGACATCCGATGCTTGTAGGCAATTTTCACTCATGTTACTCTC 1447
Db 8962 TACCTCAGATCATTTGAACGACATCCGATGCTTGTAGGCAATTTTCACTCATGTTACTCTC 9021
Qy 1448 CAGGTGAGATCAATAGGTGGCTTCTATGCTCAGGAGCTTGGGCTACACCCCTTGCAG 1507
Db 9022 CAGGTGAGATCAATAGGTGGCTTCTATGCTCAGGAGCTTGGGCTACACCCCTTGCAG 9081
Qy 1508 TCTGAGACATCGGGCCAGAGTGTCCGCTGCTAGTCTGCTCAGGAGCGGAGGCGC 1567
Db 9082 TCTGAGACATCGGGCCAGAGTGTCCGCTGCTAGTCTGCTCAGGAGCGGAGGCGC 9141
Qy 1568 CAAATTGTGGCAAGTACCTCTTCAACTGGGAGTGAAGCAACGCTTAAATCTCACTCAA 1627
Db 9142 CCAATTGTGGCAAGTACCTCTTCAACTGGGAGTGAAGCAACGCTTAAATCTCACTCAA 9201
Qy 1628 TTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1687
Db 9202 TTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9261
Qy 1688 ACATATATCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1747

Db 9262 ACATTTATCACAGCTGTCTGCTGCCGACCCCGTGTGTTTATGTTGCTGCTACTCTAC 9321
Qy 1748 TCTCGTAGGGTAGGCATCTATCTACTTCCCAACCG 1784
Db 9322 TTTCTGTAGGGTAGGCATCTACTGCTTCCCAACCG 9358
RESULT 13
HPCUNKCDS
LOCUS Hepatitis C virus mRNA, complete cds. 9414 bp linear VRL 26-JUL-1993
DEFINITION M96362.1 GI:306286
ACCESSION M96362.1
VERSION M96362.1
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES: ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1 (bases 1 to 9414)
AUTHORS Cho, J.-M.
TITLE Hepatitis C virus
JOURNAL Unpublished (1992)
COMMENT Original source text: Hepatitis C virus cDNA to mRNA.
FEATURES
location/Qualifiers
1..9414
/organism="Hepatitis C virus"
/mol_type="mRNA"
/db_xref="taxon:11103"
<1..9414
/evidence=experimental
343..9375
/note="putative"
/codon_start=1
/protein_id="AA45721.1"
/db_xref="GI:306286"
/translation="MSTNPKPQRKTRNTNRRPQDIKPPGGQIIVGGVYLLPRRGRPL
SRATKTSRSQPRGRQPIPKARPEGRWAQGPWPLVNEGGLWAGMLLSPRG
SVSMGPTDPRKSRNLGKVIDTLTCGFADLMGYIPLVGLPGGVARALAGVRVLED
GNYATCNLPCCSFSIFLLALLSLTTPVSAYEVNRNASGMVHTNDCNSSIYVEAD
MHTPCVPCVREDNSRCWALTPLAARNASVPTTLRRHVDLLVGLVAAFCASMY
VGLDGSVFVLSQFTFSPRRHETVDCNCSYVGRVSRHMRMDMMNNSPTTLLV
SOLLRIQAVVMDVTGSHGILAGLAYYSVMGNKVLIALMLFAGVDGTTHTVGGAG
GSAASLTSLSGVPVHQLQINTNGSMHNRALSNDLSLNTGFAALFKYFNAS
GCPERLATCPIDTFAGGMPITVTEPHDLDORPYOMHYAQPQCGIVPTLQVCCPVNC
FTSPVAGCTDRFAGTYRKGANETDVLNNAGPQGNWFGCTWNGGCTTCCG
PPNIGVGNNTLCTPTDPRKHPATYTKCGSGPMLTLCPLVDYVRLWHYPTCVNF
TIFKRMVYVGAERHDAACNWTGRGECDEDRSLLPLLSLTTTSMWLPSCFTTL
PALSTGLIHLQNIQVIOYLYGIVSAVSPAIKWEYIVLLPLLLADARVACALWMLL
VQAERALENLVNLNAASVAGAHGILSFVFCAAWYIKGRVPGAYALYGVWPLLL
LLALPRAYAMDRMAASCGAVFVGLVLLTSLSPHYKVFARPIWLOYLITRTEAH
LQWVPPNLNVRGRDAIILLTGVHPELIIFDIKLLAIFGLPLMVLQAGITRVPYFVR
AQLIRACMLARKVGVHYVQVMFKLAALAGTYVYDHLTPLRDMATGLRDLAVAV
PVVFSMDETKVIWAGDTAACGDIILALPASARSGKEILIGPADSLRGQWRLLAPIIT
AYSQOTRGLAGCIIISLTGRDKNOVEGVOVSTATOSFIATCINGCVTWVFGAGSK
TLAGKGPITQNTVNDQDLVGNPAPGASLTPTCTGSSDLVLTTHRADVIPIVRRRG
DGRGLLPPRPVSLKSGSGPLCPGSHAVGILPAACTRGVAMAVEFIPIVSMET
MRSPFTDNPSPAPVPTQVAHLHAPTSKSTRPAAAYAGQGYKVLNPSVAATL
GFCAYNSKAHIGIDENLTGVTITTGAPITSTYTGKFLADGGGGAGYDIIMCDECHS
TDSYTVIGITGLDQABTAGARLVLTATPGSVTVPHNLBEVALSNTGEIIPYVK
AIPIEATKGRHLIFCHSKKCDLAAKLSGLGNVAYYRGLDVSIVPTSGDVVVA
TDALMTGTFGDSVDCNTCTVQTFDLPDTFTTITTTVPQDASRSRRRTGRG
RAGIYRVTPEGRSPGMSVLCEDYDAGCAWYELTFAETSRLRAYLTPGLPVC
DHLEFSEGVFTGLTHDAHLSQKAGENPYLVAQATVCAQAAPPSPDMWRMC
LIRLPTLHGTPLLYLGLAVONEVTLTHPTKPIWTCMSADLVSTWLVGGVLA
ALAAVCTTGSVVIIGRIILSGKPAIIPDEVLVQFDEMEECASHLPYEQGMOLAE
OFQKALGLLQATKQNEAARPVESKWRALETFWAKHMFISIGIYLAGLSTLPGN
PAIRSPWATTSITISLUTHTLNLGVAQAQAPPSAASAVGAGAAVGTI
GKGLVLDIILAGYAGVAGALVAFKMSGEMPSAEDMNLPLLSLPGALVAVCAA
ILRRHVGPGEVQVMNRHIFASRGNHVSHPYVPESEPAARTQVLSLSTLTQLIK
RLHWINEDCSTPCSSSLREIWDVICTVLTDFKWLQSKLPLPGVPTFCORGKY
GVNRGDIHMTTCPCGAIITGVHNGSMRIVGPKTCNTWYGPPIINAYTGPCTPS
APNYSKALLWRVAEEVEYVTRVGDHFYVTGNTDNCPCQCPAPEFTVDGVRLR
YAPRLEPRLVEEYVQVGLHQYLSQLPCEBPBDVAULTSMULTDPSHTAETAKRLR

ARGSPPSLASSASQSLSPSLKATCTCTHHSDSPADLIEANLLWFQEMQGNITRVESN
KVIILDFDPLURAEADSGEISVPABEILRKSRKPPALPIWAPDPNPPLLESWKDDY
VPPVHSCPLPPTKAPPIPPPRKRRTVLTJESTVSSALAEALATKTGSSGSSAIDSGT
ATAPDOASGDREDSEVESPSNPPIEGEPGDDLSDGWSVTSVEBASDEVVCCSMS
YTTGALITPCAAEESKLPIPLNSLIRHHNMVYATTSRSAGLRQKRVTFDRLOVLD
DHYRDLKEMKAKASTYKALLSVEEACKLTTPHSAKSEKFGYAKDVRSLSSEKALYD
RSVMKDLLEDTETPISTTINAKNEVFCVQPEKGRKPARLIVPDLGVRVCENKALYD
VVSITLPAVMGSSYGFQYSPQRVEPLVNTWKSKKCPMGFSYDTRCFDSVTITENLRV
EESIYQCCDLAPEAKLAKSLTERLYLIGPLTNSKQNCQYRRCRASGVLTSCGNUL
TCYLKATACRAAKLRDCTMLVNGDULVICESAGTQEDDAASURVFEAMRTSAPPG
DPPOEYDLELITSCSNVSVAHDSAGRYKLYTRDPTTPLARAANETAHRTFVNSWL
GNIIMYAPTLMARMILMTHFFSILLAQEOLEKTLDCIYGACYSIEPLDLPQIIEIHL
FLSAFSLHSYSPGEINRVASCLRLKGVPLRAWRHRARSVRKLLSGRGAATCGKYL
GNMAVRTKLLKLTPIPAASRLDLSGFWVAGYSGGDIYHLSLRANRPWFMCLCLLLLSGVV
GIYLLPNR"
complement (9414)

polyA_site
ORIGIN

Query Match		89.2%;	Score 1607.4;	DB 14;	Length 9414;		
Best Local Similarity		94.0%;	Pred. No. 0;				
Matches 1671;		Conservative	0;	Mismatches	106;	Indels	0; Gaps 0;
Qy	8	GCATGTCAATGTCTTATATACATGGACAGGCGCCCTGATCACACCGCTGGCGTCCGGAGGAAA	67				
Db	7595	GCTGTTCCGATGCTACACATGGACAGGCGCCCTGATCACGCCATGCGCTCGCGAGGAAA	7654				
Qy	68	GCAAGTCCCATCAACCGCTGAGCAACTCTCTTGCTGCTGCTACACATAAAGTCTGCTATT	127				
Db	7655	GCAAGTTCGCCATCAACCGCTGAGCAATCTCTTGCTACGTCACCAACAATGCTCTATG	7714				
Qy	128	CCACAAATCCCGCAGTGCAGAGCTTCGGCAGAGAGAGGTTCACCTTTGACAGACTGCAAG	187				
Db	7715	CTACAACATCCCGCAGCGCAGSCCTTCGGCAGAGAGAGGTTCACCTTTGACAGACTGCAAG	7774				
Qy	188	TCCTTGAGGATCATTTACCGGACGCTGCTCAAGGAGATGAAGCGGAGAGCGCTCCACAGTGA	247				
Db	7775	TCCTTGAGGACACACTACCGGACGCTGCTTAAGGAGATGAAGCGGAGAGCGCTCCACAGTTA	7834				
Qy	248	AGGCTAAACTGCTATCTGTAGAAGAAGCATCAAGCTGACAGCCCGCCGCTTCGGCCAAAT	307				
Db	7835	AGGCTAAACTTCTATCTGTAGAAGAAGCTGCAAACTGACGCGCCCGCCACATTCGGSCAAAT	7894				
Qy	308	CCAAATTTGGCTATGGGCGCAAGAGCGTCCGAGGCTTATCCAGCAGGCGCGTTTAAACACA	367				
Db	7895	CCAAATTTGGCTATGGGCGCAAGAGCGTCCGAGGCTTATCCAGCAGGCGCGTTTACCCACA	7954				
Qy	368	TCCGCTCCGCTGGAGACCTTGCTGGAGGACCTGACACACCAATTCAGACACCACTCA	427				
Db	7955	TCCGCTCCGCTGGAGAGACCTTGCTGGAGGACCTGAAACCAATTAGCACTACCATCA	8014				
Qy	428	TGGCAAAAATGAGGTTTCTGCGTCCAAACAGAGAAAGAGGCGCGCAACAGCTCGCC	487				
Db	8015	TGGCAAAAATGAGGTTTCTGCTGCCAAACAGAGAGAGGCGCGCAAGCAGCTCGCC	8074				
Qy	488	TCATCGTATCCCAGACTGGGAGTTCGTGTATCGGAGAGATGGCCCTTTTACGACGTGG	547				
Db	8075	TTATCGTGTCCCAGATCTGGGAGTTCGTGTATCGGAGAGATGGCCCTTTTATCAGCTGG	8134				
Qy	548	TTTCACTCTTCTCAGCGCGTGATGGGCTCCTCATACGGATTCCAAATCTCTCCTTAAGC	607				
Db	8135	TCTCCACCTTCTCAGGCGCGTGATGGGCTCCTCATACGGATTCCAGTACTCTCCTTAAGC	8194				
Qy	608	AGCGGCTCGAGTTCCGTGGTAATACCTCGGAAGCAAGAAATGCCCTATGGGCTTCTCAT	667				
Db	8195	AGCGGCTCGAGTTCTGTGGTAATACCTCGGAATCAAGAAATGCCCTATGGGCTTCTCAT	8254				
Qy	668	ATGACACCCGCTGTTTGTGACTCAACGGTCACTTGAGAAATGACATCCGTGTTGAGGAGTCAA	727				
Db	8255	ATGACACCCGCTGTTTGTGACTCAACGGTCACTTGAGAAATGACATCCGTGTTGAGGAGTCAA	8314				
Qy	728	TTTACCAATGTTGTGACTTTGGCCCGGAGAGTACAGACGGCCATTAAGGTTCGCTCAGAGC	787				
Db	8315	TTTACCAATGTTGTGACTTTGGCCCGGAGAGCAACTGGCCATAAAGTTCGCTCAGAGC	8374				

RESULT 14
HCU16362
LOCUS

HCU16362 9415 bp RNA linear VRL 10-NOV-1994

Qy	788	GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGGCAGAACTGCGGCTATCGCGCGT	847
Db	8375	GGCTCTATATCGGGGTCCCCTGACTAATTCANAAGGGCAGAACTGCGGTTACCGCGGT	8434
Qy	848	GCGCGCGAGCGCGTGTGACGACTAGCTGGGGTAATAACCTTCACATGCTACTTTGAAGG	907
Db	8435	GCGCGCGAGCGCGTGTGACGACTAGCTGCGGTAAATACCCCTCACAATGTACCTGAAAG	8494
Qy	908	CGCTCCAGCCTGTCCAGCTGCCAGCTCCAGACTCCAGACTGACGATGCTGCTGAATGAGAGC	967
Db	8495	CCACTCGGCGCTGTCCAGCTCGGAAGCTCCGGGACTGCAACGATGCTCGTGAACGGAGAC	8554
Qy	968	ACCTTCTCGTTATCTGTGAAGCGGGGAAACCAAGAGCAACCGCAAGCTTACCGAGTCT	1027
Db	8555	ACCTTCTCGTTATCTGTGAAGCGGGGAAACCAAGAGGATGCGGCGAGCCTACGAGTCT	8614
Qy	1028	TCACGAGGCTATGACTAGGTACTCTGCCCCCCCCCTGGGGACCCGCCCAACCGGAATACG	1087
Db	8615	TCACGAGGCTATGACTAGGTACTCTGCCCCCCCCCTGGGGACCCGCCCTCAACCGGAATACG	8674
Qy	1088	ACTTGAGCTGATAACATCGTGTTCCTCCAATGTGTGCTGCTGCGCACAGATGATCTGGCA	1147
Db	8675	ACTTGAGTGTGATAACATCATGTTCCTCCAATGTGTGCTGCGCACAGATGATCTGGTA	8734
Qy	1148	AAAGGCTGACTACTCACCGGTGACCCCAACCTGCCCCCTGGCGGGCTGGTGGGAGA	1207
Db	8735	AAAGGCTGACTACTCACCGGTGACCCCAACCTGCCCCCTGGCGGGCTGGTGGGAGA	8794
Qy	1208	CAGCTAGGCACACTCCAGTCAACTCTCTGCTAGGCAACATCATGATGCGCCCACTT	1267
Db	8795	CAGCTAGACACACTCCAGTCAACTCTCTGCTAGGCAACATCATGATGCGCCCACTT	8854
Qy	1268	TGTGGCAAGGATGATTTCTGATGACTCACTTCTCTCCATCTCTTAGCCATGAGCACTT	1327
Db	8855	TATGGCAAGGATGATTTCTGATGACTCACTTCTCTCCATCTCTTAGCTCAGGAGCAAC	8914
Qy	1328	TTGAAAAGCCCTGGATTTGTGAGATCTACGGGCTTGTTCATCTCCATGAGCACTTGACC	1387
Db	8915	TTGAAAAGCCCTTAGATTTGTGAGATCTACGGGCTTGTTCATCTCCATCTCTTAGCTC	8974
Qy	1388	TACCTCAGATCAATGAACTCCAGTCTTACGCTTCTAGGCAACATCATGATGCTCTC	1447
Db	8975	TACCTCAGATCAATGAGCACTCCATGGTCTTAGCGCAATTTTCACTCCATAGTACTCTC	9034
Qy	1448	CAGCTCAGATCAATAGGCTGGCTTCATGCTCAGGAAGCTTGGGGTACCACTTCCGAG	1507
Db	9035	CAGCGAGATCAATAGGCTGGCTTCATGCTCAGAAAATCTTGGGGTACCACTTCCGAG	9094
Qy	1508	TCTGGAGACATCGGCGCAGAGTGTCCGCGCTAAAGTTACTGTCAGGGGGGAGGGCGG	1567
Db	9095	CCTGGAGACATCGGCGCAGAGTGTCCGCGCTAAGCTACTGCTCCAGGGGGGAGGGCGG	9154
Qy	1568	CAATTTGTGGCAAGTACTCTTCAACTGGGCAGTAGGACCAAGCTTAACTCACTCCAA	1627
Db	9155	CACTTTGTGGCAAGTACTCTTCAACTGGGCAGTAGGACCAAGCTCAAACTCACTCCAA	9214
Qy	1628	TTCCGCTCGCTCCCGCTGCACTTCTGCGGCTGGTTCGTTGCTGCTACAGCGGGGAG	1687
Db	9215	TCCACGCGCGTCCCCTGGTGGACTTGTCCGCGCTGGTTCGTTGCTGTTACAGCGGGGAG	9274
Qy	1688	ACATATATCACAGCTGTCTCGTCCCGACCCCGCTGGTTCAATGTTGTGCTACTCCTAC	1747
Db	9275	ACATATATCACAGCTGTCTCGTCCCGACCCCGCTGGTTCAATGTTGTGCTACTCCTAC	9334
Qy	1748	TCTCCGTAGGGGTAGGATCTATCTACTCCCCAACCG	1784
Db	9335	TTTCCGTGGGGGTAGGATCTACTCTCTCCCCAACCG	9371

DEFINITION Hepatitis C virus polyprotein gene, complete cds complete genome.

ACCESSION U16362.1 GI:567059

VERSION 1

KEYWORDS Hepatitis C virus

SOURCE Hepatitis C virus

ORGANISM Viruses; asRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1 (bases 1 to 9415)

CHO, J., Park, Y., Lee, Y., Yang, J. and Ryu, W.

Direct Submission

Submitted (27-OCT-1994) Joong Myung Cho, Lucky Biotech Research

Institute, 104-1 Munji-dong, Yoseung-ku, Daeseon 305-380, Korea

in Korea

Unpublished

2 (bases 1 to 9415)

CHO, J.

Journal Submission

Submitted (27-OCT-1994) Joong Myung Cho, Lucky Biotech Research

Institute, 104-1 Munji-dong, Yoseung-ku, Daeseon 305-380, Korea

Location/Qualifiers

1. .9415

/organism="Hepatitis C virus"

/mol_type="genomic RNA"

/db_xref="taxon:11103"

343..9375

/codon_start=1

/product="polyprotein"

/protein_id="AA52748.1"

/db_xref="GI:567060"

/translation="MSTNPKPQKTKYNNRRPQDIKPPGGQIVGGVLLPRRQPLR

GVATRKTSERQPRQPIPKARRPEGRMAQFGYPMPLVGNLEGLWAGWLLSPRG

SRPAGTDPKRSKRLGIDITLCGFADLMGYIPLVGAFLGGVARALAHGVLEID

GVNATGNLPGCSFSIFLLALLSCLTTPVSAVEVRNAGMWHVTHDNCSSNIIYEAAD

MTMHPPGVCPCRENSRSCWALTPTLAARNASVPTTTLRHVDLLVGAARFCASMY

VGDLCGSFVLSQLFTFSPRRHETVQDNCNCSYIPGRVGRHAWMDMMNNPTTALVY

SOLRLIPQAVDMFVQSHWGLIAGLAYSVMGNWAKVLIAMLLFAGVDTGTHVTCGAQ

GRAASLSLSPGVPVHQLINTNGSHINRITALSNDLSLNTGFVAALFYKRFNAS

GCPELATCRPDIPTAQMGPIYTPHDLDRPYCHVAPOPCGIYPTLVQVCGVGG

FTFSPVAVGTDRFGATYRWGANTDVLNNAGPPOGNNFCTWNGTGTFTKTCGG

PCNIGVGNMNTLCPTDGFRAHPGATYTKCSGFWLTPRCLVDYPIRLWHPCTVNF

TIFFKVMVCGAHRIDACNWRIGRCDLDRSEUSPLLLSTTEWQVLPSCFTTL

PALSTGLIHQNIIVIOYLYIGISAVVSFAIKWEYIYLLFLLADARVCAALWMLLL

VAQAALLENVLMNAASVAGAHGISFTIFFCAAWIYKGRVPGAAVALYGVWPLLL

LILLALPPRAYAMDEMAASCGAFVGLVLLTSLPHYKVFLEAFIWLQYLITRTEAH

LOVVPPLNVRGRDAIILLTCVHPHEDITKYLLALFGLPLWLAGITREVPYFVR

AGLIRACHMLARVGGHYVQVFMKLAALAGTYIYDHLTPLRDWAHTGLKDLAVVE

PVFSMETKVIWAGDTAAACGDIILACPASARRKEIILGDPADLSLEGQWRLLAIP

AYSQTRGLGCIITSLTRDNQVEGEVQVSTATQSFPLATCINGVCWTFPHGASK

TLAPKGPITQWTVNDVQDLVGPAPPGARSITPCTCGSSDLYLVTRADYIPVRRG

DGRGILLPPRPVSYLKGSGGPLLCPGSHAVGILPAVCTRGVMAVEFIIPVSMETT

MPSVPTDNPPAPVQTPQVAHLHAPTGSKSTRVPAAYAAQGVKVLNPSVAATL

GFGAHMSKHGIDPMLRTGVRTLITGAPLITYSTGKFLADGGSGGAYDIIMCDECHS

TDSTTYIGTGLVDAEAGARLVVLSTAPPGSVTPVPLNIEEVALNTGEIIPPYGK

AIPBAIKGRHLIFCHSKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVVVA

TDALMTGTFDPSVIDCNCVTQVDFSLDPTFTIETITVPODAVSRRRGTRG

RAGIYRFTGCRPSGMPDPSVLCEDYDAGCAWELTDAETSVRLRAYLNTPLGVQC

DHLFSEGVFTGLTHIDAHFLSQTKOAGENFYVAYQATVCARQAAPPSPDENWRC

LIRLKEPTHTGVPPLRLVAGVQNEVTLTHPIKFTMCMSADLEVTVSWLVGSLA

ALAAVCLITGSPVIVGRIILSGKPAIIPDREVLVQFDEMEECASHLPYFQGMOLAE

QPKQALGLOATQAAAPVESKWALETFWAKMNNFISGIOYLAGLSTLPGN

PAIRSMALFASITPLTQHTLLFNILGGMVAAQLAPPASAFVAGIAGAVGTI

GLGKVLVDLAGYAGAGVALVAFKIMSGMPSAEDVNNLLPILSPGALVGVICAA

LIRRHVPEGVQWNNRLIAPASRGNVSPRHYVPESEPAARVQILSSLTITOLK

RLHWINCEDSTPCSSWRLREIWDICTVLDFTKWLQSKLLPLRPJPFPSCKRGK

GVWRGIMHITPCCAQITGHVKNMSMRIYVGPCTCSNWTYGTPIINATYTGCTPSP

APNYSKALWAAVEEVTRVDFHYVTGMTDNVCKPCOVPAPEPEFTEVDVGLHR

YAPACRPLREBVFQVGLHQYLVGSQPLPCBEPDVAVLTSMLDPSHITAEAKRL

ARGSPSLASSASQSLAPSLKATCTTHHDSPLADLIENALLWQEMGNITRVESEN

KVYILDSPLDAEEDGELISVPAEILKRSKFPFALPIWAPNDYNNPLLESWKDPDY

VEPVHGGCLPPTKAPPIPPPRKRRTVLTTESTVSSALAEATKTFGSGSADISGT

ATAPVDAQSGDREDSDVESFSMPLEGECDPLSDGSKSTVSEBSEDVVCSSMS

YTWTCALITPCAAESKLPINPLSNLIRHNMVYATTSRAGLRQKVTYFDRLOVLD

DHYRDLKEMKAKASTVAKLUSVEACKLTPPSAKSKFGTGAKDVRSLSSRAVTHI

RSVWDLLEMTETPITSTIMAKNEVFCVQPEKGRKPARLIVFPDLGVRCVRCALYD

ORIGIN

Query Match	89.2%	Score	1607.4	DB	14	Length	9415
Best Local Similarity	94.0%	Pred. No.	0				
Matches	1671	Conservative	0	Mismatches	106	Indels	0
Gaps	0						
QY	8	GCATGCTCAATGTCCTATACATGGACGCGCCCTGATCACACCGTGGCTCGGAGGAAA	67				
DB	7595	GCTGTTCATGTCCTACACATGGACGCGCCCTGATCACCGCATGCCCTCGGAGGAAA	7654				
QY	68	GCAAGCTGCCATCAACCGCTGAGCAACTCTTCTGCTGCTCACCATAACTTGTCTATT	127				
DB	7655	GCAAGTTGCCATCAACCGCTTGAAGCAATCTTCTGCTAGCTCACCAACATGGTCTATG	7714				
QY	128	CCACAAATCCCGAGTCCAAAGCCCTGCGGAGAGAGAGGTCACTTTGACAGACTGCAAG	187				
DB	7715	CTACAAATCCCGAGTCCAAAGCCCTGCGGAGAGAGGTCACTTTGACAGACTGCAAG	7774				
QY	188	TCCTGGACGATCATTACCGGACGCTGCTCAAGGAGATCAAGCGGAGCGTCCACAGTGA	247				
DB	7775	TCCTGGACGACCATACCGGACGCTGCTTAAAGGAGATGAAGCGGAGCGTCCACAGTTA	7834				
QY	248	AGGCTAAACTGCTATCTCTAGAAGAGCATGCAAGCTGACGCGCCCGCATTCGGGCAAT	307				
DB	7835	AGGCTAAACTGCTATCTCTAGAAGAGCATGCAAGCTGACGCGCCCGCATTCGGGCAAT	7894				
QY	308	CCAAATTTGGCTATGGGCAAGAGGACGCTCGGAGGCTATCCAGCAGGCGCGTTTACCCACA	367				
DB	7895	CCAAATTTGGCTATGGGCAAGAGGACGCTCGGAGGCTATCCAGCAGGCGCGTTTACCCACA	7954				
QY	368	TCGCTCGCTGGAGGACGCTGCTGGAGGACACTGACACACCAATTCAGACACCATCA	427				
DB	7955	TCGCTCGCTGGAGGACGCTGCTGGAGGACACTGACACACCAATTCAGACACCATCA	8014				
QY	428	TGGCAAAATAGAGTTCGTCGCTCCAAACAGAGAGGAGGCGCGCAACCGCTCGCC	487				
DB	8015	TGGCAAAATAGAGTTCGTCGCTCCAAACAGAGAGGAGGCGCGCAACCGCTCGCC	8074				
QY	488	TCATCGTATTCACAGACCTGGGAGTTCGTGTATGCGAGAGATGCCCTTTACGACGTGG	547				
DB	8075	TTATCGTATTCACAGACCTGGGAGTTCGTGTATGCGAGAGATGCCCTTTACGACGTGG	8134				
QY	548	TTTCCACTCTTCACGCGCTGATGGGCTCTCATACGGAATCCAATCTCTCTAAGC	607				
DB	8135	TCTCCACCTTCTCAGGCGCTGATGGGCTCTCATACGGAATCCAATCTCTCTAAGC	8194				
QY	608	AGCGGTGAGTTCCTGCTGTAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT	667				
DB	8195	AGCGGTGAGTTCCTGCTGTAATACCTGGAAATCAAAAGAAATGCCCTATGGGCTTCTCAT	8254				
QY	668	ATGACACCGCTGTTTGTGACTCAACGCTCACTGAGAAATGACATCCGTGTTGAGGAGTCAA	727				
DB	8255	ATGACACCGCTGTTTGTGACTCAACGCTCACTGAGAAATGACATCCGTGTTGAGGAGTCAA	8314				
QY	728	TTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCCATAGGTCGCTCACAGAGC	787				
DB	8315	TTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCCATAGGTCGCTCACAGAGC	8374				
QY	788	GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGCAGAACTGGGGCTATCCCGGT	847				
DB	8375	GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGCAGAACTGGGGCTATCCCGGT	8434				
QY	848	CCCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAATACCTCACATGCTACTTTGAAGG	907				
DB	8435	CCCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAATACCTCACATGCTACTTTGAAGG	8494				

Qy	908	CCGCTGCAGCCTGTCGAGCTGCCAAGCTCCAGGACTCCAGGACTCGATGCTCGTGAATGGAGACG	967
Ds	8495	CCACTGCGGCTGTCGAGCTCGAGCTCCGGAGCTCCGGAGCTGCAGATGCTCGTGAACGGAGACG	8554
Qy	968	ACCTTGCTGTTATCTGTGAAGCGCGGAAACCCAGAGAGCGCGGCAAGCTTACGAGTCT	1027
Ds	8555	ACCTTGCTGTTATCTGTGAAGCGCGGAAACCCAGAGAGTGCAGGAGCTTACGAGTCT	8614
Qy	1028	TCACGAGGCTATGACTAGTACTCTGCGCCCTCCGCGGAGCCGCGCCCAACCGGATACG	1087
Ds	8615	TCACGAGGCTATGACTAGTACTCTGCGCCCTCCGCGGAGCCGCGCTCAACCGGATACG	8674
Qy	1088	ACTTGGAGCTCATACATCGTGTCTCCCAATGTCGTCGACACGATGTCATCTGGCA	1147
Ds	8675	ACTTGGAGTTCATACATCATGTTCTCCCAATGTCGTCGACACGATGTCATCTGGTA	8734
Qy	1148	AAAGGCTGACTACTCCTCACCCGTCACCCACCGTCCCTTTCGCGGCGCTGCGTGGGAGA	1207
Ds	8735	AAAGGCTGACTACTCCTCACCCGTCACCCACCGTCCCTTTCGACGGGCTGCGTGGGAGA	8794
Qy	1208	CAGCTAGGACACTCCAGTCAACTCCTGGTAGGCAACATCATCATGATGCGCCCACTT	1267
Ds	8795	CAGCTAGGACACTCCAGTCAACTCCTGGTAGGCAACATCATCATGATGCGCCCACT	8854
Qy	1268	TGTGGGCAAGGATGATTCGATGACTCATCTTCTCTCATCTTCTAGCCAGGAGCAAC	1327
Ds	8855	TATGGGCAAGGATGATTCGATGACTCATCTTCTCTCATCTTCTAGCCAGGAGCAAC	8914
Qy	1328	TGGAAGGCGCTGATGTCAGATCTACGGGCTGTGTACTCCATGAGCCACTTCAC	1387
Ds	8915	TGGAAGGCGCTGATGTCAGATCTACGGGCTGTGTACTCCATGAGCCACTTCATC	8974
Qy	1388	TACCTCAGATCATTAAGCACTCCATGCTTTAGCGCATTTTCACTCCATGATTACTCTC	1447
Ds	8975	TACCTCAGATCATTAAGGAGCTCCATGCTTTAGCGCATTTTCACTCCATGATTACTCTC	9034
Qy	1448	CAGGTGAGATCAATAGGTGCTTTCATGCTCAGGAAAGCTTGGGGTACCACTTGGCGAG	1507
Ds	9035	CAGGCGAGATCAATAGGTGCTTTCATGCTCAGGAAAGCTTGGGGTACCACTTGGCGAG	9094
Qy	1508	TCTGGAGACATCGGCGCAGAGTGTCCGCTAGTACTGTCCTCCAGGGGGGAGGCGG	1567
Ds	9095	CCTGGAGACATCGGCGCAGAGTGTCCGCTAGTACTGTCCTCCAGGGGGGAGGCGG	9154
Qy	1568	CCATTGTGGCAAGTACCTTTCAACTGGGCGAGTAAGGACCAAGCTTAACTCACTCAA	1627
Ds	9155	CCATTGTGGCAAGTACCTTTCAACTGGGCGGTAAGGACCAAGCTTAACTCACTCAA	9214
Qy	1628	TTCCGGCTCGCTCCGGGCTGGACTTGTCCGGCTGGTTCGTTGCTGGCTACAGCGGGGAG	1687
Ds	9215	TCCAGCGCGTCCGGGTTGGACTTGTCCGGCTGGTTCGTTGCTGGTACAGCGGGGAG	9274
Qy	1688	ACATATATCACAGCTGTCTGTCGCGGACCCGCTGTTTCAATGTTGCTTACTCTTAC	1747
Ds	9275	ACATATATCACAGCTGTCTGTCGCGGACCCGCTGTTTCAATGTTGCTTACTCTTAC	9334
Qy	1748	TCTCGTAGGGGTAGGCATCTATCTACTTCCCAACCCG	1784
Ds	9335	TCTCGTAGGGGTAGGCATCTATCTACTTCCCAACCCG	9371
RESULT 15			
LOCUS	HPCJ483	9448 bp	RNA linear VRL 01-FEB-2000
DEFINITION	Hepatitis C virus genome, complete sequence.		
ACCESSION	D13558	D01217	
VERSION	D13558.1	GI:221604	
KEYWORDS	C protein; E protein; E2 protein; NS1 protein; NS2 protein; NS3 protein; NS4 protein; NS5 protein.		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		

Hepacivirus.
1 (bases 1 to 9448)
Okamoto,H., Kojima,M., Okada,S., Yoshizawa,H., Iizuka,H., Tanaka,T., Muchmore,E.E., Peterson,D.A., Ito,Y. and Mishiro,S.
Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability
Virology 190 (2), 894-899 (1992)
92391112
1325713
2 (bases 1 to 9448)
Okamoto,H.
Direct Submission
Submitted (17-OCT-1991) Hiroaki Okamoto, Jichi Medical School, Immunology Division: Minamikawachi-machi, Kawachi-gun, Tochigi 329-04, Japan (E-mail:hokamoto@jichi.ac.jp,
Tel:0285-44-2111(ex.3334), Fax:0285-44-1357)
Submitted (17-Oct-1991) to DDBJ by:
Hiroaki Okamoto
Immunology Division
Jichi Medical School
Kawachi-gun
Tochigi 329-04
Japan
Phone: 0285-44-2111 x3334
Fax: 0285-44-1557.
Location/Qualifiers
1. .9448
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HC-J4"
/isolate="HC-J4/83"
/db_xref="taxon:11103"
342. .9374
/codon_start=1
/product="polyprotein precursor"
/protein_id="BAA02756.1"
/db_xref="GI:221605"
/translation="MSTNPKPKYKRNTRRPQDVKPGGQIVGGVYLLPRGRPL
GVRATKTSERSQPRGWRQPIKPAIRPEGRAWAQPGYPMPLVGNELGAGWMLSPRG
SRPSGPTDPRRRSRNLKVIDTLCGTFADLMGYIPLVAGIPLGAARALAHGVRVLE
GVNATLNPGRSFSIFLLALLSLTIPASAVEVRNVSIGIYHVTDCNSSTIVYEAAD
MIMHTPCGVCVREONSRCWALPTTLAARNASPTTIRRHVDLLVCAAPFCASMY
VLDLGSFVLVSQLFTFSPRRHTVQDCNSIYPGHSLSHRNANDMMWNSPTTALVY
SOLLIPQAVVDMKAGHWGLAGIAYSMVGNWAKVLIIVALLFAGDGETTSGGA
SHTTSLASLSPGASQRIQLVNTNGSWHINRTALNCDSLHTGFLAALFYTHRFNS
GCPERMASCRPIDWFAQMGPIITYTEPDSPDQRPYCHYAPRPGTVPASOVCGPVY
FSPSPVVGCTDRSGVPTYSWGENETDMLNNTRPDGNWFCCTWMNSTGFTKTCGG
PFCNIGGVNHTLTCTDCFRKPEATYTKCGSGPWLTPRCLVDYPRLWHPCTFNF
SIFKVMYVGGVEHRLNAACTWTRGERCNLEDRSELSPLLLSTTEWQILPCAFTLL
PALSTGLIHQNIYDVQYLVGVSFAVSVFAIKWEYILLLFLLADARVACACMLMLL
IAQAAALNVLVINAASVAGAGHGLVGLVFLTSLPYKVFRLRLIMLQVFTIRAEAH
LILALPPRAYAMDREMAACGGAVLGLVFLTSLPYKVFRLRLIMLQVFTIRAEAH
MQWVPELVNVRGDAIILLTCAVHPELIPDITKLALLPLMLVLAQITVPPVPR
AQGLIRACMLRKVAGGHTVQWAPKLGALTGTYYVNHILTPLDWHAHAGRLAVAE
PVFSMDTKVITWAGDAAACQDIIILGLPVSARRKEIFLGPADSLEGGQWELLAPIT
AYVQSTRGVLGCIITSLTRGDKNQVEGVVSTATQSFATCINGSCWTVYHGAGSK
TLIAGPKGPIQWTVNLDLGVQAOAPPGARSMTPTCTGSSDLVYLRHADVIPRRRG
DSRGSLSPRVSYLKGSSGGLLCPSGHVGVVRAAVCTRGVAKAVDFIPVESMETI
MRSPTFDNSSPPAVPTQVLAHLHAPTSCGSKTKVPAAYAGQYKAVLVNLSVAATL
GFGAYMPKAGHDPIRTGRTITITGGPITYTYGKFLADGGCGSGGAYDIIICDECHS
TDSITTLIGITVLDQATAGARLVLAATATPPGSTVTPHNIEEIGLSNNGELIPFYGK
AIPDAIKGRHILFCHSKKCDLAALKTGLNNAVAYYRGLDVSVIPPIGDVAIVA
TDALMTGFTGDSVIDCNTCTVDFSLDPTFTIETTTVQDQAVSRQRRGTGRG
RSGIYBFTVPGSPSCFDSVLCVCAQAGWVLTPAETSVRLAVATNPLGLPVCO
DHLEFWSFTGLSHDAHLPSQTKAGDNFPYLVAYQNTVCARAQAPPPSNDQWKK
LIRLKPHTLHGPTLLYRGLAGQNEVTLTPTIKYINACMSADLEVSTISWLVVGVVLA
ALAAAYCLTTGSSVIVGRIILSGKPAVPDPREVLVYQFDEMEBCASQLPPIEQMQLA
QFKQALGLIQTAKOAAAPVWESKRALETFWAKHWNFTISGQYLAGLSTLPGN
PAFASLMAFTASITSLPTTQNTLLFNILGWWAAQLAPPASAFVAGAGAAVAGSI
GIGKVLVDILLAGGAGVAGLVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVGVA
ILRRHVPGEGAGVQWNRLLIAFASRGNHVSHPHYVESDAARAVTQILSSLTITOLLK
ILRHWINEBCSTPCSGWLRDVMWDVICTVITDFKTLWLSKLPRLPGVPLFSCORGYK
GLWVGDIQMTCTPCCGAGHGVKNGSMRIVGPRTCNSNTWHGTFPINAITYTGCTPSP

FEATURES

source

CDS

```

APNYSALNRVRAAEYVEVTRVGDPHYVTGMTNDNVKCPQVPAPEFFTEVDVGRILHR
IYAPACPLLRDVAFOGLNQVLVGSQPCPEPDVTVLTSMITDPSHITATKRRIL
AGSPSLASASOLASPSLAKATCTTHDSPADLIEANLLRWEMGNI TRVSEN
KVVILDSFPLAEGEREISVAABETLRKSRKPSALPTWAPDYNPPLLESWKDPDY
VPPVHGCPPLPTKAPPDIPBRKRTVLVTESNVSSALAEALATKTFSSGSAVDUST
ATALPDQASDDGDKGSDVESYSMPLEGPDPDLSDGWSWSTVEEASEDVCOSMS
YTHGALITPCAEBSKLPINLSNLSLRHNMVYATTSRSASLRQKVTFFDLQVLD
DHYRDLKEMKAKSTVRKLLSIEACKLTPHSAKSKFGYKADVRNLSSRAVNH
RSVWDLLEDTEPTDITIMAKNEVFCVQEPKRPARELI VPPDLGVRCHEMALYD
VVSITPQAVMGSYGFQYSPKORVPLANTWKSKECPMGFSVDTRCFSTVTESDIRV
ESIVOCCLDAPARQOARSLTERLYIGGLPLNSKONGCYRRCASGLVITTSQNTL
TCYLKATACRAAKLODCTMLVNGDDLVVICESAGTOEDAAALRVFTAMTRYSAAPP
DPPQEDYDELITSCSSNVSAHDAGKRVYVLTDRPTIPLARAAMETARTPPVNSPL
GNIIMYAPALMARMLTMTFFSILLAEQLEALDCOYGVACYSIEPLDLPOILBERLH
GLSAPLSHSYSGEINRVASCLURKLGVPLRVHRASVRKLLSQGRATCGRYL
PNMAVRTKLKLPPIPAASQLDLSCHFPVAGYSGDLYHLSLRARPRWFLLCLLUSVG
GLVLLPNR"
mat_peptide 342..914
/product="C protein"
mat_peptide 915..1490
/product="E protein"
mat_peptide 1491..2528
/product="NS1/E2 protein"
mat_peptide 2529..3359
/product="NS2 protein"
mat_peptide 3360..5186
/product="NS3 protein"
mat_peptide 5187..6380
/product="NS4 protein"
mat_peptide 6381..9371
/product="NS5 protein"
stem_loop 6005..6104
/note="possible stem-loop structure in NS4 sequence"
misc_feature 9413..9448
/note="T-stretch of 36 bp"

ORIGIN
Query Match 89.2%; Score 1607.4; DB 14; Length 9448;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 8 GCATGCTCAATGCTCTATACATGAGCAGCGCCCTGATCACACCGTGGCGTGGAGGANA 67
Db |||||
7594 GCTGCTCAATGCTCTATACGTTGGACAGCGCCCTGATCACGCCATGGCGTGGAGGANA 7653

Qy 68 GCATGCTGCCATCAACCGCGTGGAGCAACTCTTGTGCGTCAACATAACCTGCTATT 127
Db |||||
7654 GTAAAGTGGCCATCAACCCGTTGAGCAACTCTTGTGCGTCAACACATGCTGTACG 7713

Qy 128 CCACAACATCCGCGAGTGCAGGCTGCGGCGAGAGAGGTCACTTTTGACAGACTGCAAG 187
Db |||||
7714 CCACAACATCCGCGAGCGCAAGCTCCGCGAGAGAGGTCACTTTTGACAGACTGCAAG 7773

Qy 188 TCCTGGACGATCATTTACGGGACTGCTCAAGGAGATGAAGCGAAGCGGTCCACAGTGA 247
Db |||||
7774 TCCTGGACGATCATTTACGGGAGCTGCTCAAGGAGATGAAGCGAAGCGGTCCACAGTGA 7833

Qy 248 AGGCTAAACTGCTATCTGTAGAAGACATGCAAGCTGACGCGCCCGCATTCGCGCAAT 307
Db |||||
7834 AGGCTAAACTGCTATCTATAGAAGAGCGCTGCAAGCTGACGCGCCCGCATTCGCGCAAT 7893

Qy 308 CCAATTTGGCTATGGGGCAAGGACGTCGGAGGACCTTCCAGCAGGCGCGTTAACACA 367
Db |||||
7894 CCAATTTGGCTATGGGGCAAGGACGTCGGAGGACCTTCCAGCAGGCGCGTTAACACA 7953

Qy 368 TCCGCTCCGTGTGAAGGACTTGTGTGAGGACACTGACACACCAATTCAGACCAACCATCA 427
Db |||||
7954 TCCGCTCCGTGTGGAGGACTTGTGTGAGGACACTGACACACCAATTCAGACCAACCATCA 8013

Qy 428 TGGCAAAATAGAGTTTCTGCGTCCAAACAGAGAAAGAGCGCCGGAACCAAGCTCGCC 487
Db |||||
8014 TGGCAAAATAGAGTTTCTGCGTCCAAACAGAGAAAGAGCGCCGGAACCAAGCTCGCC 8073

```

```

498 TCATCGTATTCACACACCTGGAGTTCGTGTATGGAGAGATGSCCCTTTACGACGTGG 547
Db |||||
8074 TTATCGTATTCACACACCTGGAGTTCGTGTATGGAGAGATGSCCCTTTACGACGTGG 8133

Qy 548 TTTCCACTCTTCTCAGCGCGGTGATGGGCTCTCTCATACGGATTCCTCAATCTCTCTAAGC 607
Db |||||
8134 TCTCCACCTTCTCAGCGCGGTGATGGGCTCTCTCATACGGATTCCTCAATCTCTCTAAGC 8193

Qy 608 AGCGGTCGAGTTCCTGCTGTAATACCTGGAAAGCAAGAAATGCCCTATGGCTTCTCAT 667
Db |||||
8194 AGCGGTCGAGTTCCTGCTGTAATACCTGGAAATCAAGAAATGCCCTATGGCTTCTCAT 8253

Qy 668 ATGACACCGCTGTTTGAATCAACGGTCACTGAGAAATGACATCCGTTTGAAGAGTCAA 727
Db |||||
8254 ATGACACCGCTGTTTGAATCAACGGTCACTGAGAGTGACATTCGTGTTGAGGAGTCAA 8313

Qy 728 TTTACCAATGTTGACTTTGGCCCCCGAAGCTAGACAGGCGCATAGGTGCGTCAAGAGC 787
Db |||||
8314 TTTACCAATGTTGACTTTGGCCCCCGAAGCTAGACAGGCGCATAGGTGCGTCAAGAGC 8373

Qy 788 GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGCGGCTATCGCGGT 847
Db |||||
8374 GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGCGGCTATCGCGGT 8433

Qy 848 GCGCGCAGCGCGCTGCTGACGACTAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 907
Db |||||
8434 GCGCGCAGCGCGCTGCTGACGACTAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8493

Qy 908 CCGCTGACGCTGCTGAGCTGCGCAAGCTCCAGGACTCCAGATGCTGCTGCTGCTGCTGCTG 967
Db |||||
8494 CCGCTGACGCTGCTGAGCTGCGCAAGCTCCAGGACTCCAGATGCTGCTGCTGCTGCTGCTG 8553

Qy 968 ACCTGTGCTGTTATCTGTGAAAGCGCGGAAACCCCAAGAGGACGCGGCAAGCTACGACT 1027
Db |||||
8554 ACCTGTGCTGTTATCTGTGAAAGCGCGGAAACCCCAAGAGGATGCGGCGGCTACGACT 8613

Qy 1028 TCAGGAGGCTATGACTAGGTACTCTGCGCCCTCTGGGACCCCGCCCAACCGGAAATACG 1087
Db |||||
8614 TCAGGAGGCTATGACTAGGTACTCTGCGCCCTCTGGGACCCCGGGAATCGCGCCCAACCG 8673

Qy 1088 ACTTGAGCTGATTAACATCTGTTCTCTCAATGTGTGGTGGCGACAGATGCTGCTGCTG 1147
Db |||||
8674 ACTTGAGCTGATTAACATCTGTTCTCTCAACCGTGCAGTCGCGCACGATGCTGCTGCTG 8733

Qy 1148 AAAGGTGTACTACTCTACCGGTGACCCACCGTCCCGCTTGGCGGGCTGCTGCTGCTGCTG 1207
Db |||||
8734 AAAGGTGTACTACTCTACCGGTGACCCACCGTCCCGCTTGGCGGGCTGCTGCTGCTGCTG 8793

Qy 1208 CAGCTAGGCACACTCCAGTCAACTCTGCTGCTAGGCAACATCATGATGCGGCCACTT 1267
Db |||||
8794 CAGCTAGGCACACTCCAGTCAACTCTGCTGCTAGGCAATCATCATGATGCGGCCACTT 8853

Qy 1268 TGTGGCAAGGATGATCTCTGACTCAGTCTCTTCTCCATCTCTTAGCCCGAGGCAAC 1327
Db |||||
8854 TGTGGCAAGGATGATCTCTGACTCAGTCTCTTCTCCATCTCTTAGCTCAAGAGCAAC 8913

Qy 1328 TTGAAAAGCCCTGGATGTTGATGATCTAGCGGCTTGTGTACTCCATTTGAGCCACTTGAAC 1387
Db |||||
8914 TTGAAAAGCCCTGGATGTTGATGATCTAGCGGCTTGTGTACTCCATTTGAGCCACTTGAAC 8973

Qy 1388 TACCTCAGATCATTTGAACGACTCCATGCTTCTAGCGCAATTTTACCTCCTAGTACTCTC 1447
Db |||||
8974 TACCTCAGATCATTTGAACGACTCCATGCTTCTAGCGCAATTTTACCTCCTAGTACTCTC 9033

Qy 1448 CAGGTGAGATCAATAGGCTGGCTTCTAGCTCAGGAAGCTGCGGTAACCACTTCGCGAG 1507
Db |||||
9034 CAGGTGAGATCAATAGGCTGGCTTCTAGCTCAGGAAGCTGCGGTAACCACTTCGCGAG 9093

Qy 1508 TCTGAGACATCGGCGCAGAGTGTCCGCGCTTAAGTTACTGTCCAGGGGGGAGGCGCG 1567
Db |||||
9094 TCTGAGACATCGGCGCAGAGTGTCCGCGCTTAAGTTACTGTCCAGGGGGGAGGCGCG 9153

Qy 1568 CCAATTTGGCAAGTACCTCTTCAACTGGGCGAGTAAGGACCAAGCTTAACTCACTCCAA 1627

```

Db	9154	CCACTTGTGGCAGATACCTCTTTAACTGGGCAGTAGGACCAAGCTTAAACTCACTCCAA	9213
Qy	1628	TTCCGGCTGCGTCCCGGCTGGACTTGTCCGGCTGGTTCCGTTGCTGGCTACAGCGGGGAG	1687
Db	9214	TCCCGGCTGCGTCCCGAGCTGGACTTGTCCGGCTGGTTCCGCTGGTTACAGCGGGGAG	9273
Qy	1688	ACATATATCACAGCCTGTCTCGTGCCTCCGACCCCGCTGGTTTCANCTTGTGCTACTCCTAC	1747
Db	9274	ACATATATCACAGCCTGTCTCGTGCCTCCGACCCCGCTGGTTTCTGTGTGCTACTCCTAC	9333
Qy	1748	TCTCGTAGGGGTAGGCATCTATCTACTCCCAACCG	1784
Db	9334	TTTCTGTAGGGTAGGCATTTACCTGTCTCCCAACCG	9370

Search completed: October 3, 2005, 18:55:29
Job time : 7975.72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 10:33:35 ; Search time 6149.94 Seconds
(without alignments)
11159.430 Million cell updates/sec

Title: US-10-712-479-1
Perfect score: 1803
Sequence: 1 atggctagcatgtcaatgc.....ggcatcaccatcaccatcac 1803

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gsel: *
9: gb_gsel2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.4	2.5	891	1	AL516365
2	40.8	2.3	873	5	BQ938432
3	40.8	2.3	6636	3	BC043482
4	40.6	2.3	593	6	CA390175
5	40.6	2.3	665	7	CA423579
6	39	2.2	487	5	BX282027
7	39	2.2	685	6	CD579994
8	39	2.2	820	4	BG818891
9	39	2.2	851	7	CK192072
10	39	2.2	882	5	BQ898223
11	39	2.2	938	5	BQ720502
12	39	2.2	941	5	BQ720548
13	39	2.2	954	5	BQ673878
14	39	2.2	993	2	BE263117
15	39	2.2	1606	3	CR620795
16	39	2.2	1737	3	CR591176
17	39	2.2	1739	3	CR605251
18	39	2.2	1740	3	CR615736
19	39	2.2	1742	3	CR605475
20	39	2.2	1747	3	CR598826
21	39	2.2	1753	3	CR624274
22	39	2.2	1758	3	CR609048
23	39	2.2	1763	3	CR624218
24	39	2.2	1769	3	CR596536

25	39	2.2	1812	3	CR609614
26	39	2.2	1828	3	CR610243
27	38.6	2.1	604	6	CD216287
c 28	38.4	2.1	288	2	BB388175
c 29	38.4	2.1	1071	8	AF094871
c 30	38	2.1	403	8	AO936657
31	38	2.1	509	1	AI657258
32	38	2.1	554	8	CC094862
33	37.8	2.1	612	2	BE497710
c 34	37.8	2.1	772	9	CNS02XEP
c 35	37.8	2.1	869	9	CNS03W08
c 36	37.8	2.1	1012	9	CNS01SJJK
c 37	37.6	2.1	700	5	BU448257
c 38	37.4	2.1	259	8	AZ865259
39	37.4	2.1	546	2	AW255326
40	37.4	2.1	905	9	AG042203
41	37.2	2.1	841	1	AL516952
42	37.2	2.1	958	4	BT828644
43	37.2	2.1	1007	5	BQ672629
44	37.2	2.1	1100	9	CNS014RI
c 45	37	2.1	274	9	CE434162

ALIGNMENTS

RESULT 1
AL516365
LOCUS
DEFINITION
AL516365 Homo sapiens NEUROBLASTOMA Homo sapiens cdna clone
CS0DA005YJ06 5-PRIME, mRNA sequence.
ACCESSION
AL516365
VERSION
AL516365.3 GI:45653002
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 891)
AUTHORS
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
On Feb 13, 2001 this sequence version replaced gi:30490924.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 6432.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0DA005DE03QP16c=6432.r.

FEATURES

source
Location/Qualifiers
1..891
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA005YJ06"
/cissue_type="NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 2.5%; Score 45.4; DB 1; Length 891;
Best Local Similarity 47.8%; Pred. No. 0.19;
Matches 118; Conservative 3; Mismatches 126; Indels 0; Gaps 0;

QY 37 GCCCTGATCACACCGTGGCTGGCGGAGAAAGCAAGCTGCCATCAACGGCTTGAGCAAC 96
 DB 582 CGCGCGCTCAACACGAGTTGGAGGAGGAGACCGCAAGGNGGACGGGASACCAAC 641
 QY 97 TCCCTGCTGCTCACCATAACCTGGTCTATTCCACACATCCCGCAGTGAAGCTCGGG 156
 DB 642 CTCGTGCTCTTCCGCAAGGAGCTGACGATCCCACTCTGTCGCCCTGGAACTAGAGCGC 701
 QY 157 CAGAGAGAGGTCACCTTTGACAGACTGCAAGTCTCTGGACGATCAATACCGGAGCTGCTC 216
 DB 702 AAGATGAGTCTCTGATGATGAGATTGAGTTCTCTCAGAGCTGCACGAGGAGAGCTG 761
 QY 217 AAGAGATGAAGGCGGAGGCGCTCCACAGTGAAGCTAACTGCTATCTGTAGAAGACCA 276
 DB 762 CGAGAMCTGCGAGTGCAGTGTGGAGAGCCAGCAGTGGCAGCAGGTGGAGGTGGAAGCCACG 821
 QY 277 TGCAAGC 283
 DB 822 GKGAAGC 828

RESULT 2
 BQ938432 873 bp mRNA linear EST 21-AUG-2002
 LOCUS AGENCOURT_8924837 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6467897
 DEFINITION 5', mRNA sequence.

ACCESSION BQ938432
 VERSION BQ938432.1 GI:22353910
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL13994 row: i column: 18
 High quality sequence stop: 709.
 Location/Qualifiers

FEATURES
 source
 1..873
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6467897"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 94"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
 Site2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 2.3%; Score 40.8; DB 5; Length 873;
 Best Local Similarity 48.3%; Pred. No. 4.3;
 Matches 114; Conservative 0; Mismatches 122; Indels 0; Gaps 0;
 QY 309 CAAATTTGGCTATGGGCAAGGACGTCGGAGCCTATCCAGCAGGCGCGTTAACCAACAT 368
 DB 192 CCATCTTTGGGAGAGTTTCAGAGGACAGCTCCCGCATATCATCATCTCTTTCGCGCTCTT 251

QY 369 CGCTCCGTGTGCAAGGACTGTCTGGAGGACACTGCACACCAATTCAGACCAACCATCAT 428
 DB 252 CCAGTCATGAGCTGTGTCAAGCTTCTGAGTGAAGGTGAGGGATCCGCACACTGCTCTG 311
 QY 429 GCGAAATAATGAGGTTTTCTGGCTCCACAGAGAAAGAGGCGCGAAACAGAGCTCGCT 488
 DB 312 GACATTCATCAAGTCTTCCAGGCTTGCCTATGTGGCACTTCTCATAGCAATGATATT 371
 QY 489 CATCGTATTCGCCAGACTGGGAGTTGCTGTATCGGAGAGATGGCCCTTTAGCAAG 544
 DB 372 CTTCACTATGCAAGTCATTGGCATGCAGATGTTTGGCAAGGTGGCTCTTCAGAGCG 427

RESULT 3
 BC043482 6636 bp mRNA linear HTC 19-NOV-2003
 LOCUS Mus musculus calcium channel, voltage-dependent, alpha 1F subunit,
 DEFINITION mRNA (cDNA clone IMAGE:5369391), containing frame-shift errors.
 ACCESSION BC043482.1 GI:28175783
 VERSION BC043482
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 6636)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Loquellano,N.A., Peters,G.J.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., R.A.,
 Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smailus,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 6636)
 Strausberg,R.
 Direct Submission
 Submitted (09-JAN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saedi, Jacqueline
 Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Query Match 2.3%; Score 40.6; DB 7; Length 665;
Best Local Similarity 50.8%; Pred. No. 4.5;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 60 GGAGGAAGCAAGCTGCCATCAACGGCTGAGCAACTCTTGTGGTCACTAACCT 119
DB 350 GTACCAAAACAGGGTGGCATGAACCTGGGCAGACCTGCTTCTCTCCCGACATCCT 409
QY 120 GGTCTATTCCACACATCCCGCAGTGAAGCTCGGCAGAGAAGTCACTTTGACAG 179
DB 410 CCTCCACACAGCAATAGCGAAGAGTACAACTTTCTGTAGATGAAGCTATGACCAAGAA 469
QY 180 ACTGCAAGTCTCGACGATCATACCGGAGCTGCTCAAGCAGATGAAGCGGAGGGTCT 239
DB 470 ATGCCATGTCCGTCGCCACAGCAGGATGTATTGTCACAGATGAATTAGAGAGAG 529
QY 240 CACAGTGAAGG 250
DB 530 GAAGATGAACG 540

RESULT 6

LOCUS BX282027 487 bp mRNA linear EST 04-MAR-2003
DEFINITION BX282027 NCI CGAP Brn67 Homo sapiens cDNA clone IMAGE998E2010822 ;
IMAGE:4914811, mRNA sequence.
ACCESSION BX282027
VERSION BX282027.1 GI:28612877
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 487)
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.
TITLE Human Unigeneset - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE998E2010822.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13u, Primer sequence: CCGTGTAAACGACGGCCAGT.

FEATURES

source
1..487
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE998E2010822 ; IMAGE:4914811"
/tissue_type="anaplastic oligodendroglioma with lp/19q

loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP_Brn67"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 2.2%; Score 39; DB 5; Length 487;
Best Local Similarity 47.4%; Pred. No. 12;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
QY 37 GCCTCATCACACCGTGCCTCGGAGGAAAGCAAGCTGCCCATCAACCGCTGAGCAAC 96
DB 11 GCGGCGCTCAAGCAGAGGTTGGAGGAGGAGACGCGCAAGCGGAGGACGCGGAGCAAC 70
QY 97 TCCTTGTGGTCAACATTAACCTGGTCTATTCCACAAATCCCGCAGTCAAGCCCTGCCG 156
DB 71 CTGCTGCTCTTCGCAAGGACGTGGACGATGCCACTCTGTCCGCGCTGGAATCTAGAGCG 130
QY 157 CACAAGAAGTCACTTTGACAGACTGCAAGTCTCGACGATCATACCGGAGCTGAGCTCTC 216
DB 131 AAGATTGAGTCTCTGATGGATGAGATGATTTCTCAAGAGTGCACGAGGAGGAGCTG 190
QY 217 AAGGAGATGAAGCGAAGCGCTCCACAGTGAAGGCTAAACTGCTATCTGTAGAGAAGCA 276
DB 191 CGAGACTTCAGTGATGTGGAGAGCCAGGTCGACAGTGGAGTGGAAAGCCACG 250
QY 277 TGCAAGC 283
DB 251 GTGAAGC 257

RESULT 7

LOCUS CD579994
DEFINITION CD579994 685 bp mRNA linear EST 18-MAY-2004
EST PSF001449 hex1 (RZPD no. 800) Homo sapiens cDNA clone
MPMGp800L12569 5' similar to PERIPHERIN.
[Source:SWISSPROT;Acc:P41219], mRNA sequence.
ACCESSION CD579994
VERSION CD579994.2 GI:47495907
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 685)
AUTHORS Bussow,K., Cahill,D., Nietfeld,W., Bancroft,D., Scherzinger,E.,
Lehrach,H. and Walter,G.
TITLE A method for global protein expression and antibody screening on
high-density filters of an arrayed cDNA library
Nucleic Acids Res. 26 (21), 5007-5008 (1998)
JOURNAL 98451599
MEDLINE 9776767
PUBMED
COMMENT On Jun 13, 2003 this sequence version replaced gi:31744319.
Contact: Bussow K
PSF, http://www.proteinstrukturfabrik.de
Max Planck Institute of Molecular Genetics
Innestr. 73, 14195 Berlin, Germany
Tel: +49 30 32639 2802
Fax: +49 30 32639 2833
Email: bussow@molgen.mpg.de
PSF clone ID = 13446. Clone expresses a soluble His-tag fusion
protein of 45 kDa size (good expression).
Seq primer: pQ65, TGAGCGGATAACAATTTACACAG
High quality sequence stop: 411.
Location/Qualifiers
1..685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MPMGp800L12569"

/tissue_type="brain"
 /dev_stage="embryonic, 14.8 and 15.8 weeks (determined by foot length)"
 /clone_lib="HEX1 (RZPD no. 800)"
 /note="Vector: pQE30NST (AF074376); Site 1: SalI; Site 2: NotI; 1st strand cDNA was prepared from mRNA obtained from brain material of two feti, 14.8 and 15.8 weeks old, with a Not I - oligo(dT) primer [5', GACTAGTTCCTAGATCGGAGCGCGCTTTTTTTTTTT 3']
 Double-stranded cDNA was ligated into Sal I adaptors, digested with Not I and cloned into the Not I and Sal I sites of the pQE30NST vector. This cDNA expression library represents a subset of clones that were selected for protein expression and has been normalised by oligonucleotide fingerprinting. The library was constructed by Konrad Bussow."

ORIGIN

Query Match 2.2%; Score 39; DB 6; Length 685;
 Best Local Similarity 47.4%; Pred. No. 13;
 Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCCTGATCACCGTGGCGTCCGAGGAAAGCAAGCTGCCCATCAACGCGCTGAGCAAC 96
 |||
 Db 36 GCGCGCTCAAGCAGAGGTTGGAGGAGGAGACGCGCAAGCGGAGGCGCGGAGCACAAC 95

QY 97 TCCTTGCTGCTCACCATAACTGCTCTATTCCACAACATCCGCGAGTGCAGCCTGCGG 156
 |||
 Db 96 CTCGTGCTCTTCCGCAAGGAGCTGGACGATGCACTCTGTCGCGCTGGAACCTAGAGCGC 155

QY 157 CAGAAGAAGGTCCACCTTTGACAGAGCTGCAAGTCTCGGACGATCATTTACCGGAGCTGCTC 216
 |||
 Db 156 AGATTGAGTCTCTGATGATGAGATTGAGTTCTTCAAGAAGCTGCACGAGGAGGAGCTG 215

QY 217 AAGGAGATGAAGCGGAGCGTCCACAGTGAAGGCTAAACTGCTATCTGTAGAAGAAGCA 276
 |||
 Db 216 CGAGAGCTGCAGGTGAGTGTGGAGAGCAGGAGGTGCAGAGCTGGAGGTGGAAGCCACG 275

QY 277 TGCAGC 283
 |||
 Db 276 GTGAAGC 282

RESULT 8
 BG818891
 LOCUS 60279275P2 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4914811
 DEFINITION 5', mRNA sequence.
 ACCESSION BG818891
 VERSION BG818891.1 GI:14166478
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 820)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LAM10822 row: e column: 20
 High quality sequence stop: 783.
 Location/Qualifiers
 1..820
 /organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4914811"
 /tissue_type="anaplastic oligodendroglioma with lp/19q loss"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Brn67"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 2.2%; Score 39; DB 4; Length 820;
 Best Local Similarity 47.4%; Pred. No. 14;
 Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCCTGATCACACCGTGGCGGAGGAAAGCAAGCTGCCCATCAACGCGCTGAGCAAC 96
 |||
 Db 11 GCGCGCTCAAGCAGAGGTTGGAGGAGGAGACGCGCAAGCGGAGGAGCGCGGAGCACAAC 70

QY 97 TCCTTGCTGCTCACCATAACTGCTCTATTCCACAACATCCCGCAGTGCAGGCTGCGG 156
 |||
 Db 71 CTCGTGCTCTTCCGCAAGGAGCTGGACGATGCCACTCTGTCCCGCTGGAACCTAGAGCGC 130

QY 157 CAGAAGAAGGTCCACCTTTGACAGACTGCAAGTCTCGGACGATCATTTACCGGAGCTGCTC 216
 |||
 Db 131 AAGATTGAGTCTCTGATGATGAGATTGAGTTCTTCAAGAAGCTGCACGAGGAGGAGCTG 190

QY 217 AAGGAGATGAAGCGGAGCGTCCACAGTGAAGGCTAAACTGCTATCTGTAGAAGAAGCA 276
 |||
 Db 191 CGAGAGCTGCAGGTGAGTGTGGAGAGCAGCAGGTGCAGGAGGTGGAGGTGGAAGCCACG 250

QY 277 TGCAGC 283
 |||
 Db 251 GTGAAGC 257

RESULT 9

CK192072
 LOCUS EST781387 BEA Boophilus microplus cDNA clone BEAD742, mRNA
 DEFINITION 851 bp
 ACCESSION CK192072
 VERSION CK192072.1 GI:49572606
 KEYWORDS EST.
 SOURCE Boophilus microplus (southern cattle tick)
 ORGANISM Boophilus microplus
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari; Parasitiformes; Ixodida; Ixodoidea; Ixodidae; Boophilus.
 REFERENCE 1 (bases 1 to 851)
 AUTHORS Nene,V., Quackenbush,J., George,J. and Guerrero,F.
 TITLE An index of genes transcribed in the tick Boophilus microplus
 JOURNAL Unpublished (2004)
 COMMENT Other ESTs: EST781386
 Contact: Vishvanath Nene
 Parasite Genomics
 The Institute For Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-610-5968
 Fax: 301-838-0208
 Email: nene@tigr.org
 Seq primer: M13 reverse.
 Location/Qualifiers
 1..851
 /organism="Boophilus microplus"
 /mol_type="mRNA"
 /db_xref="taxon:6941"
 /clone="BEAD742"
 /dev_stage="Adult and larvae"
 /lab_host="E. coli strain DH10B-TonA"
 /clone_lib="BEA"
 /note="Organ: Whole ticks and dissected organs; Vector:

pExpress 1: A normalized cDNA library was custom built by Express Genomics. Oligo-dT primed cDNA was directionally cloned into NotI-EcoRV site of pExpress 1. Universal M13 primers were used to generate 5' and 3' EST data."

ORIGIN

Query Match 2.2%; Score 39; DB 7; Length 851;
Best Local Similarity 48.4%; Pred. No. 14;
Matches 108; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 1003 GAGGACGGCGAAGCCTACGAGTCTTCCACGGAGGCTATGACTAGTACTCTGCCCCCCT 1062
DB 392 GAGGACCTGGAGACCTACGACTGTTCTCGCGTGTGCTGCTGAGTACTGCGTGGCCTG 451

QY 1063 GGGGACCGCCCGAACCGGAATACGACTTGGAGCTGATACATCGTGTTCCTCCAAATGTG 1122
DB 452 GATGAGCAGGAGGTAGCGGAACAGGCATTAACTCATCTTGGCGTTTGACGAGATTGTG 511

QY 1123 TCGGTGCGACAGTATCTGGCAAAAGGGTGTACTACCTCACCCTGACCCACCGCTC 1182
DB 512 GCCCTGGGCTACCGGAGAGCGTCAACCTGGCTCAATCCGACCTTCGTTGAAATGAC 571

QY 1183 CCCCTTGGCGGCTGCGTGGGAGACAGCTAGGCACACTCCAG 1225
DB 572 TCGCATGAGGAGAGGTCTACCGAGCTGTTGACAGACACAG 614

RESULT 10

BQ898223
LOCUS BQ898223
DEFINITION AGENCOURT_8045553 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6212523
5', mRNA sequence.

ACCESSION BQ898223

VERSION BQ898223.1 GI:22290237

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 882)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2376 row: m column: 04

High quality sequence stop: 588.

Location/Qualifiers

1. 882

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6212523"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_102"

/note="Organ: salivary gland; Vector: pORF7; Site_1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed

by Ling Hong in the laboratory of Gerald M. Rubin

(University of California, Berkeley) using ZAP-cDNA

synthesis kit (Stratagene) and Superscript II RT (Life

Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity 47.4%; Pred. No. 15;

Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCTCATCACACCGTCGCTGCGGAGGAAGCAAGCTGCCCATCAACGCGCTGAGCAAC 96

DB 272 GCGCGCTCAAGCAGAGGTTGGAGGAGGAGACCGCAAGCGGAGGACCGGAGCAAC 331

QY 97 TCCTTCGTCGTCACCATAACTGGTCTATTCCACAACATCCCGCAGTCGAAGCCCTGCCG 156

DB 332 CTGCTGCTCTTCGCGANGACGTGGACGATGCCACTCTGTCCGCTGGAATCTAGAGCGC 391

QY 157 CAGAAGAAGTCACCTTTGACAGACTGCAAGTCTCTGGACGATCATTTACCGGAGCTGCTC 216

DB 392 AAGATTGAGTCTCTGATGGATGAGATTGAGTTCTTCAAGAAGCTGCACGAGGAGGCTG 451

QY 217 AAGGAGATGAAGCGAGCGCTCCACAGTGAAGGCTAAACTGCTATCTGTAGAGGAAGCA 276

DB 452 CGAGACCTGCAGGTGAGTGTGGAGAGCCAGGTCGACGAGGTGGAGGTGGAGGCCACG 511

QY 277 TGCAGGC 283

DB 512 GTGAAGC 518

RESULT 11

BQ720502

LOCUS BQ720502

DEFINITION AGENCOURT_8219550 Lupski sympathetic_trunk Homo sapiens cDNA clone

IMAGE:6188094 5', mRNA sequence.

ACCESSION BQ720502

VERSION BQ720502.1 GI:21859399

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 938)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LHAM3583 row: c column: 07

High quality sequence stop: 627.

Location/Qualifiers

1. 938

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6188094"

/sex="male"

/tissue_type="sympathetic trunk"

/dev_stage="adult, 16 yr"

/lab_host="DH10B"

/clone_lib="Lupski sympathetic trunk"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:

NotI; Site_2: SalI; cDNA made by oligo-dT priming.

Directionally cloned using the following adaptors:

5'-TCGACCCAGTCGTCGCGCG-3' and

5'-GACTTACTTACGATCGGAGCGCGCCCTT(15)-3'. Size selected >

1 kb for average insert length 1.9 kb. This is a primary

library, non-amplified. Library constructed by Life

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylour

College of Medicine); available through Life

Technologies."

```

ORIGIN
Query Match          2.2%; Score 39; DB 5; Length 938;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCTGATCACACCGTGCCTCGGAGGAAAGCAAGTGCCTCAATCAACGCGCTGAGCAAC 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 GCGGCGCTCAAGCAGAGGTTGGAGGAGGAGACGCGCAAGCGGGAGGACGCGGAGCACAAC 289

QY 97 TCCTTGTCTCGCTCACCAATACCTGCTATTTCACAAATCCCGCAGTGCAGCCCTGCGG 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 CTCGTGCTCTCCGAGAGACGTGGACGATGCACTCTGTCCGCTCGAACTAGAGCGC 349

QY 157 CAGAAGAAGGTCAACCTTTGACAGACTGCAAGTCTCTGGAGCATCAATACCGGAGCTGCTC 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 AAGATTGAGTCTCTGATGATGAGATTGAGTTCTCTCAAGAGCTCCAGAGAGGAGCTG 409

QY 217 AAGGAGATGAAGCGGAGCGTCCACAGTGAAGGCTTAACTGCTATCTGTAGAAGCA 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 CGAGACCTGCAGGTGAGTGTGGAGAGCCAGCAGGTGCAGCAGGTGGAGGTGGAGCCACG 469

QY 277 TGCAGC 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 GTGAAGC 476

RESULT 12
BQ720548
LOCUS BQ720548
DEFINITION AGENCOURT 8219731 Lupski synepathetic_trunk Homo sapiens cDNA clone
IMAGE:6189155 5', mRNA sequence.
ACCESSION BQ720548
VERSION BQ720548.1 GI:21859445
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13583 row: e column: 20
High quality sequence stop: 555.
Location/Qualifiers
1. .941
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6189155"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski synepathetic trunk"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCCACCCACCGGTCG-3' and
5'-GACTAGTCTAGATCGGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor

```

```

ORIGIN
Query Match          2.2%; Score 39; DB 5; Length 941;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCTGATCACACCGTGCCTCGGAGGAAAGCAAGTGCCTCAATCAACGCGCTGAGCAAC 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 GCGGCGCTCAAGCAGAGGTTGGAGGAGGAGACGCGCAAGCGGGAGGACGCGGAGCACAAC 232

QY 97 TCCTTGTCTCGCTCACCAATACCTGCTATTTCACAAATCCCGCAGTGCAGCCCTGCGG 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 CTCGTGCTCTCCGAGAGACGTGGACGATGCCACTCTGTCCCGCTCGAACTAGAGCGC 292

QY 157 CAGAAGAAGGTCAACCTTTGACAGACTGCAAGTCTCTGGAGCATCAATACCGGAGCTGCTC 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 AAGATTGAGTCTCTGATGATGAGATTGAGTTCTCTCAAGAGCTCCAGAGAGGAGCTG 352

QY 217 AAGGAGATGAAGCGGAGCGTCCACAGTGAAGGCTTAACTGCTATCTGTAGAAGCA 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 CGAGACCTGCAGGTGAGTGTGGAGAGCCAGCAGGTGCAGCAGGTGGAGGTGGAAGCCACG 412

QY 277 TGCAGC 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 413 GTGAAGC 419

RESULT 13
BQ673878
LOCUS BQ673878
DEFINITION AGENCOURT 8035094 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6212691
5', mRNA sequence.
ACCESSION BQ673878
VERSION BQ673878.1 GI:21784724
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ruben Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2377 row: d column: 04
High quality sequence stop: 492.
Location/Qualifiers
1. .954
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6212691"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 102"
/note="Organ: salivary gland; Vector: pOT87; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

```

ORIGIN

Query Match 2.2%; Score 39; DB 5; Length 954;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCCTGATCACACCGTGGCTCGCGAGGAAAGCAAGCTGCCCATCAACGGCGTGAGCAAC 96
DB 272 CGCGCGCTCAAGCAGAGGTTGGAGSAGGAGAGCGCGAAGCGGGAGGAGCGGAGCACAAC 331

QY 97 TCCCTGCTGCTACCAATACCTGGTCTATTCCACACATATCCCGAGTGCAGCTGCTCGG 156
DB 332 CTGCTGCTCTTCCGCAAGGAGGTGGACGATCCCACTCTGTCCCGCTGGAACTAGAGCGC 391

QY 157 CAGAAGAGGTCACCTTTGACAGACTGCAAGTCTCTGGACGATCATTAACCGGAGCTGCTC 216
DB 392 AAGATTGAGTCTCTGATGGATGAGATTGAGTTCTCTCAAGAAAGCTGCACGAGGAGCTG 451

QY 217 AAGGAGATGAAGGCAAGCGCTCCACAGTGAAGGCTAAACTGTCTATCTGTAGAAGAAGCA 276
DB 452 CGAGACCTCGAGTGAGTGTGAGAGCCAGCAGGTGCAGCAGGTGGAGTGGAGTGGAGCCACG 511

QY 277 TCGAAGC 283
DB 512 GTGAAGC 518

RESULT 14

BE263117
LOCUS BE263117
DEFINITION BE263117.1 GI:9136653
ACCESSION BE263117
VERSION BE263117.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 993)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Plate: LiCM116 row: 9 column: 07
High quality sequence stop: 699.

FEATURES

source
1..993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3160350"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 19"
/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 2.2%; Score 39; DB 2; Length 993;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCCTGATCACACCGTGGCTCGCGAGGAAAGCAAGCTGCCCATCAACGGCGTGAGCAAC 96
DB 339 CGCGCGCTCAAGCAGAGGTTGGAGGAGGAGACGCGCAAGCGGGAGGAGCGGAGCACAAC 398

QY 97 TCCCTGCTGCTACCAATACCTGGTCTATTCCACACATATCCCGAGTGCAGCTGCGG 156
DB 399 CTGCTGCTCTTCCGCAAGGACGTTGGACGATGCCACTCTCTGTCGCGCTGGAACTAGAGCGC 458

QY 157 CAGAAGAGGTCACCTTTGACAGACTGCAAGTCTCTGGACGATCATTAACCGGAGCTGCTC 216
DB 459 AAGATTGAGTCTCTGATGGATGAGATTGAGTTCTCTCAAGAAAGCTGCACGAGGAGCTG 518

QY 217 AAGGAGATGAAGGCAAGCGCTCCACAGTGAAGGCTAAACTGTCTATCTGTAGAAGAAGCA 276
DB 519 CGAGACCTCGAGTGTGAGTGTGGAGCCAGCAGGTGCACGAGTGGAGTGGAGCCACG 578

QY 277 TCGAAGC 283
DB 579 GTGAAGC 585

RESULT 15

CR620795
LOCUS CR620795
DEFINITION CR620795.1 GI:50501602
ACCESSION CR620795
VERSION CR620795.1
KEYWORDS HPC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1606)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1606)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source
1..1606
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA010Y118"
/tissue_type="Neuroblastoma"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 2.2%; Score 39; DB 3; Length 1606;
Best Local Similarity 47.4%; Pred. No. 18;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 37 GCCCTGATCACACCGTGGCTCGCGAGGAAAGCAAGCTGCCCATCAACGGCGTGAGCAAC 96
DB 419 CGCGCGCTCAAGCAGAGGTTGGAGGAGGAGACGCGCAAGCGGGAGGAGCGGAGCACAAC 478

QY 97 TCCCTGCTGCTACCAATACCTGGTCTATTCCACACATCCCGAGTGCAGCTGCGG 156
DB 479 CTGCTGCTCTTCCGCAAGGACGTTGGACGATGCCACTCTGTCCCGCTGGAACTAGAGCGC 538

QY 157 CAGAAGAGGTCACCTTTGACAGACTGCAAGTCTCTGGACGATCATTAACCGGAGCTGCTC 216
DB 539 AAGATTGAGTCTCTGATGGATGAGATTGAGTTCTCTCAAGAAAGCTGCACGAGGAGCTG 598

Qy 217 AAGGAGATGAAGCGAAGCGTCCACAGTGAAGGCTAAACTGCTATCTGTAGAAGCA 276
Db 599 CGAGACCTGCAGGTGAGTGTGGAGAGCCAGCAGGTGCAGCAGGTGGAGGTGAAGCCACG 658
Qy 277 TGAAGC 283
Db 659 GTGAGC 665

Search completed: October 3, 2005, 22:17:00
Job time : 6154.94 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 09:20:55 ; Search time 994.373 Seconds
(without alignments)
10733.697 Million cell updates/sec

Title: US-10-712-479-1

Perfect score: 1803

Sequence: 1 atggtagcatgtcatgtc.....ggcatcaccatcaccatcac 1803

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1803	100.0	1803	12	ADO19020
2	1722	95.5	1740	12	ADO19022 DNA encod
3	1607.4	89.2	1884	6	ADG19089 HCV NS5B
4	1607.4	89.2	9414	13	ADR82190 Hepatitis
5	1607.4	89.2	9472	2	AAQ33282 Korean he
6	1597.8	88.6	5955	8	ACA61114 HCV cDNA
7	1597.8	88.6	9587	13	ADR82189 Hepatitis
8	1596.2	88.5	7979	10	ADD93729 Hepatitis
9	1596.2	88.5	7979	10	ADD93730 Hepatitis
10	1596.2	88.5	7979	10	ADD93732 Hepatitis
11	1596.2	88.5	7979	10	ADD93731 Hepatitis
12	1596.2	88.5	9595	2	AAQ33282 Hepatitis
13	1596.2	88.5	9595	4	AAQ33282 Hepatitis
14	1596.2	88.5	9595	4	AAQ33282 Hepatitis
15	1596.2	88.5	9595	12	ADO36222 Hepatitis
16	1596.2	88.5	9595	12	ADO36222 Hepatitis
17	1585.8	88.5	1779	10	AAQ33282 Hepatitis
18	1594.6	88.4	7989	10	ADD93733
19	1594.4	88.4	1776	10	AAQ33282 Hepatitis
20	1594.2	88.4	1778	10	AAQ33282 Hepatitis

21	1593	88.4	9416	2	AAQ20268
22	1593	88.4	9416	2	AAQ21829
23	1592.8	88.3	37090	8	ACA61113
24	1591.4	88.3	9565	8	ACA61111
25	1591.4	88.3	7989	10	ADD93728
26	1591.4	88.3	9416	2	AAQ33282
27	1591.4	88.3	9436	2	AAQ33282
28	1591	88.2	1779	2	AAQ33282
29	1591	88.2	1779	10	AAQ33282
30	1589.8	88.2	9400	13	ADR82191
31	1589.4	88.2	1779	2	AAQ33282
32	1589.4	88.2	1779	2	AAQ33282
33	1588.4	88.1	1779	2	AAQ33282
34	1586.6	88.0	3123	12	ADM29102
35	1585	87.9	7141	6	AAQ33282
36	1585	87.9	7789	6	AAQ33282
37	1585	87.9	7789	6	AAQ33282
38	1585	87.9	7789	6	AAQ33282
39	1585	87.9	7789	6	AAQ33282
40	1585	87.9	7789	6	AAQ33282
41	1585	87.9	7789	6	AAQ33282
42	1585	87.9	7789	6	AAQ33282
43	1585	87.9	7789	6	AAQ33282
44	1585	87.9	7789	6	AAQ33282
45	1585	87.9	7789	10	ADD93722

ALIGNMENTS

RESULT 1
ADO19020

ID ADO19020 standard; DNA; 1803 BP.

AC ADO19020;

DT 12-AUG-2004 (first entry)

DE DNA encoding full-length HCV NS5B polymerase, FL NS5B.

KW RNA polymerase activity; continuous-read assay;

KW Hepatitis C virus infection; HCV infection; polymerase; NS5B;

KW bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;

KW diabetes; ocular disorder; renal dysfunction; lymphoma;

KW lymphoproliferative disorder; metabolic disorder; arthritis;

KW sleep disorder; thyroid disorder; anti-inflammatory; hepatotropic;

KW viricide; osteopathic; cytostatic; antidiabetic; ophthalmological;

KW nephrotropic; antiarthritic; gene; ds.

OS Hepatitis C virus.

PH Key Location/Qualifiers

FT CDS 1..1803

FT /tag= a

FT /partial

FT /product= "FL NS5B"

FT /note= "This sequence lacks a stop codon"

XX WO2004044228-A2.

XX 27-MAY-2004.

XX 13-NOV-2003; 2003WO-US036465.

XX 13-NOV-2002; 2002US-0425981P.

XX (PFIZ) PFIZER INC.

XX Yagi Y, Sheets MP, Wells PA, Poorman RA, Epps DE;

XX Morgan AG;

XX WPI: 2004-420337/39.

XX P-PSUB; ADO19021.

XX Detecting RNA polymerase activity in a continuous-read manner, useful in
PT treating osteoporosis, carcinomas, cardiovascular diseases, ocular
PT disorders or arthritis, by contacting an RNA polymerase with an
PT oligonucleotide template.

Disclosure; SEQ ID NO 1; 46pp; English.

XX The present invention relates to a method for detecting RNA polymerase
CC activity in a continuous-read manner. The method comprises contacting an
CC RNA polymerase with an oligonucleotide template in a reaction mixture
CC comprising an assay buffer, under conditions in which the RNA polymerase
CC is active, adding a fluorescent dye capable of binding double-stranded
CC nucleic acid molecules to the reaction mixture, and measuring the
CC fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis
CC C virus (HCV) polymerase, NS5B. The method is useful for treating HCV
CC infection, bone mineral diseases like osteoporosis, carcinomas,
CC cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,
CC lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,
CC sleep disorders and thyroid disorders. The present sequence encodes full-
CC length HCV NS5B polymerase (designated FL NS5B).

XX Sequence 1803 BP; 414 A; 529 C; 471 G; 389 T; 0 U; 0 Other;

Query Match 100.0%; Score 1803; DB 12; Length 1803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTAGCATGTCAATGTCTTATACATGAGCAGCGCCCTGATCAGACCGTGGCTGGG 60
DB 1 ATGGCTAGCATGTCAATGTCTTATACATGAGCAGCGCCCTGATCAGACCGTGGCTGGG 60

QY 61 GAGGAAGCAAGCTGCCATCAACCGCTGAGCACTCCTTGGCTGCCTCACCATAACCTG 120
DB 61 GAGGAAGCAAGCTGCCATCAACCGCTGAGCACTCCTTGGCTGCCTCACCATAACCTG 120

QY 121 GTCTATTCACAAATCCCGAGTCAAGCTCGCGCAGAGAAGGTCACTTTGACAGA 180
DB 121 GTCTATTCACAAATCCCGAGTCAAGCTCGCGCAGAGAAGGTCACTTTGACAGA 180

QY 181 CTGCAAGTCTCGGAGCATCAITACCGGACGTGCTCAAGGAGATGAAGCGGAGCGTCC 240
DB 181 CTGCAAGTCTCGGAGCATCAITACCGGAGTGTCTCAAGGAGATGAAGCGGAGCGTCC 240

QY 241 ACAGTGAAGCGTAAACTGCTATCTGTAGAGAGCATGCAAGCTGACGCCCGCGCATTCG 300
DB 241 ACAGTGAAGCGTAAACTGCTATCTGTAGAGAGCATGCAAGCTGACGCCCGCGCATTCG 300

QY 301 GCCAAATCCAAATTTGGCTATGGGCGAAAGGACGTCCCGAGCCTATCCAGCAGGSCCGTT 360
DB 301 GCCAAATCCAAATTTGGCTATGGGCGAAAGGACGTCCCGAGCCTATCCAGCAGGSCCGTT 360

QY 361 AACCATCATCGGTCGGTGTGGAAGGACTTGTCTGGAGGACACTGACACACCAATTGAGACC 420
DB 361 AACCATCATCGGTCGGTGTGGAAGGACTTGTCTGGAGGACACTGACACACCAATTGAGACC 420

QY 421 ACCATCATGCGCAAAAATGAGGTTTTCTGCGTCCAAACGAGAAAAGGAGCGCGCAAAACA 480
DB 421 ACCATCATGCGCAAAAATGAGGTTTTCTGCGTCCAAACGAGAAAAGGAGCGCGCAAAACA 480

QY 481 GTCGGCTCATCGTATTCAGACCTGGGAGTTCGTGTATGAGAGAGTGGCCCTTTAC 540
DB 481 GTCGGCTCATCGTATTCAGACCTGGGAGTTCGTGTATGAGAGAGTGGCCCTTTAC 540

QY 541 GAGCTGGTTCCACTCTTCTCAGGCCGTGATGGGCTCCTCATACGATTTCCAACTACTCT 600
DB 541 GAGCTGGTTCCACTCTTCTCAGGCCGTGATGGGCTCCTCATACGATTTCCAACTACTCT 600

QY 601 CCTAAGCAGCGGTCGAGTTCTCGTGAATACCTGCAAAAGCAAGAAATGCCCTATGGGC 660
DB 601 CCTAAGCAGCGGTCGAGTTCTCGTGAATACCTGCAAAAGCAAGAAATGCCCTATGGGC 660

QY 661 TTCTCATATGACACCGCGCTGTTTGTGACTCAACCGGTCACTGAGAAATGACATCCGTTGAG 720

DB 661 TTCTCATATGACACCGCGCTGTTTGTGACTCAACCGGTCACTGAGAAATGACATCCGTTGAG 720
QY 721 GAGTCAATTTACCAATGTTGTGACTTTGGCCCCCGAGCTAGACAGGCCATTAAGTGTGCTC 780
DB 721 GAGTCAATTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCCCAAGTGTGCTC 780
QY 781 ACAGAGCGGCTCATGTCCGGGGTCCCATGACTTAACCTCAAAAGGCGAGAACTAC 840
DB 781 ACAGAGCGGCTCATGTCCGGGGTCCCATGACTTAACCTCAAAAGGCGAGAACTAC 840
QY 841 CGCCGGTCCGGCGAGCGGCTGTGACACTAGCTGCGGTAAATACCTCAGATGCTAC 900
DB 841 CGCCGGTCCGGCGAGCGGCTGTGACACTAGCTGCGGTAAATACCTCAGATGCTAC 900
QY 901 TTGAAGCGCGCTGACGCTGTCTGAGCTGCAAGCTCCAGGACTCCAGGACTGCAAGTGTCTGTAAT 960
DB 901 TTGAAGCGCGCTGACGCTGTCTGAGCTGCAAGCTCCAGGACTCCAGGACTGCTGTAAT 960
QY 961 GGAGACGACCTTCTGTTATCTGTGAAAGCGCGGGAACCCAGAGAGACGCGCAAGCCTA 1020
DB 961 GGAGACGACCTTCTGTTATCTGTGAAAGCGCGGGAACCCAGAGAGACGCGCAAGCCTA 1020
QY 1021 CGAGTCTTCACGAGGCTATGACTAGGTACTCTGCCCCCTCGGGACCCGCCCAACCG 1080
DB 1021 CGAGTCTTCACGAGGCTATGACTAGGTACTCTGCCCCCTCGGGACCCGCCCAACCG 1080
QY 1081 GAATACGACTTGGAGCTGATAACATCGTCTTCCAATGTGTCCGTCGACACGATGCA 1140
DB 1081 GAATACGACTTGGAGCTGATAACATCGTCTTCCAATGTGTCCGTCGACACGATGCA 1140
QY 1141 TCTGGCAAAAAGGGTGACTACCTCACCGGTGACCCCAACCGTCCCTTGGCGGGCTGCG 1200
DB 1141 TCTGGCAAAAAGGGTGACTACCTCACCGGTGACCCCAACCGTCCCTTGGCGGGCTGCG 1200
QY 1201 TGGGAGACAGCTAGGACACTCCAGTCAACTCTTGGCTAGGCAACATCATCATGTATGCG 1260
DB 1201 TGGGAGACAGCTAGGACACTCCAGTCAACTCTTGGCTAGGCAACATCATCATGTATGCG 1260
QY 1261 CCCACTTTGTGGCGCAGGAGTATTCTGATGACTACCTTCTTCCATCTTCTAGGCCAG 1320
DB 1261 CCCACTTTGTGGCGCAGGAGTATTCTGATGACTACCTTCTTCCATCTTCTAGGCCAG 1320
QY 1321 GAGCAACTTGAAAAAGCCCTGGAGTTGTGAGATCTAGCGGCTTTGTTACTCCATTTAGCCCA 1380
DB 1321 GAGCAACTTGAAAAAGCCCTGGAGTTGTGAGATCTAGCGGCTTTGTTACTCCATTTAGCCCA 1380
QY 1381 CTTGACCTACTCAGATCATTTGAAAGCATCCATGCTTGTAGCGCATTTTCACTTCCATAGT 1440
DB 1381 CTTGACCTACTCAGATCATTTGAAAGCATCCATGCTTGTAGCGCATTTTCACTTCCATAGT 1440
QY 1441 TACTCTCAGGTGAGATCAATAGGGTGGCTTCACTGCTCAGGAGCTTGGGTTACCAACC 1500
DB 1441 TACTCTCAGGTGAGATCAATAGGGTGGCTTCACTGCTCAGGAGCTTGGGTTACCAACC 1500
QY 1501 TTGGGAGTCTCGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGG 1560
DB 1501 TTGGGAGTCTCGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGG 1560
QY 1561 AGGGCCGCCATTTGTGGCAAGTACTCTTTCAACTTGGGCGAGTAAGGACCAAGCTTAAACTC 1620
DB 1561 AGGGCCGCCATTTGTGGCAAGTACTCTTTCAACTTGGGCGAGTAAGGACCAAGCTTAAACTC 1620
QY 1621 ACTCCAAATCCGGCTGGCTCCCGCTGGAATTTGTGTCGGCTGGTTCGTTGCTTACAGC 1680
DB 1621 ACTCCAAATCCGGCTGGCTCCCGCTGGAATTTGTGTCGGCTGGTTCGTTGCTTACAGC 1680
QY 1681 GGGGAGACATATATCAGAGCTGTCTGTCGCCGACCCCGCTGTTTCACTGTGTGCTGCTA 1740
DB 1681 GGGGAGACATATATCAGAGCTGTCTGTCGCCGACCCCGCTGTTTCACTGTGTGCTGCTA 1740
QY 1741 CTCTTACTCTCGTAGGGGTAGGCATCTATCTACTTCCCCAAACCGGCATCACCATCACCAT 1800

Db	1741	CTTCCTACTCTCCGTAGGGTAGGCATCTATCTACTCCCAACCGGCATCCACCATCACCAT	1800
Qy	1801	CAC 1803	
Db	1801	CAC 1803	
RESULT 2			
ADO19022			
ID	ADO19022	standard; DNA; 1740 BP.	
XX	AC	ADO19022;	
XX	DT	12-AUG-2004 (first entry)	
XX	DE	DNA encoding C-terminally truncated HCV NS5B polymerase, Cdelta21 NS5B.	
KW		RNA polymerase activity; continuous-read assay;	
KW		Hepatitis C virus infection; HCV infection; polymerase; NS5B;	
KW		bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;	
KW		diabetes; ocular disorder; renal dysfunction; lymphoma;	
KW		lymphoproliferative disorder; metabolic disorder; arthritis;	
KW		sleep disorder; thyroid disorder; antiinflammatory; hepatotropic;	
KW		viricide; osteopathic; cytostatic; antidiabetic; ophthalmological;	
KW		nephrotropic; antiarthritic; gene; ds.	
OS		Hepatitis C virus.	
XX			
Key		Location/Qualifiers	
FT	CDS	1..1740	
FT		/*tag= a	
FT		/partial	
FT		/product= "Cdelta21 NS5B"	
FT		/note= "This sequence lacks a stop codon"	
XX			
PN	WO200404228-A2.		
XX			
PD	27-MAY-2004.		
XX			
XX			
PF	13-NOV-2003; 2003WO-US036465.		
XX			
PR	13-NOV-2002; 2002US-0425981P.		
XX	(PFIZ) PFIZER INC.		
PA			
XX			
PI	Yagi Y, Sheets MP, Wells PA, Shelly JA, Poorman RA, Epps DE;		
PI	Morgan AG;		
XX			
DR	WPI; 2004-420337/39.		
DR	P-PSDB; ADO19023.		
XX			
PT	Detecting RNA polymerase activity in a continuous-read manner, useful in		
PT	treating osteoporosis, carcinomas, cardiovascular diseases, ocular		
PT	disorders or arthritis, by contacting an RNA polymerase with an		
PT	oligonucleotide template.		
XX			
PS	Disclosure; SEQ ID NO 3; 46pp; English.		
XX			
CC	The present invention relates to a method for detecting RNA polymerase		
CC	activity in a continuous-read manner. The method comprises contacting an		
CC	RNA polymerase with an oligonucleotide template in a reaction mixture		
CC	comprising an assay buffer, under conditions in which the RNA polymerase		
CC	is active, adding a fluorescent dye capable of binding double-stranded		
CC	nucleic acid molecules to the reaction mixture, and measuring the		
CC	fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis		
CC	C virus (HCV) polymerase, NS5B. The method is useful for treating HCV		
CC	infection, bone mineral diseases like osteoporosis, carcinomas,		
CC	cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,		
CC	lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,		
CC	sleep disorders and thyroid disorders. The present sequence encodes C-		
CC	terminally truncated HCV NS5B polymerase (designated Cdelta21 NS5B).		
XX			
SQ	Sequence 1740 BP; 404 A; 509 C; 457 G; 370 T; 0 U; 0 Other;		

Db 464 TCCGCTCCGTGTTGGAGGACTTGTCTGGAAAGACACTGTAGACACCAATTTGACACCAATCA 523
Qy 428 TGGCAAAAATAGAGTCTTCTCGCTCAACACAGAGAAAGAGCGCGCAAAACAGCTCGCC 487
Db 524 TGGCAAAAATAGAGTCTTCTCGCTCAACACAGAGAAAGAGCGCGCAAAACAGCTCGCC 583
Qy 488 TCATCGTATTCAGACCTTGGAGTTCGTGTATGCGAGAAGATGCGCCCTTACAGAGTGG 547
Db 584 TTATCGTATTCAGACCTTGGAGTTCGTGTATGCGAGAAGATGCGCCCTTACAGAGTGG 643
Qy 548 TTTCACACTCTTCTCAGGCGGTGATGGGCTCTCTCATAGCGATTCAGATCTCTCTAAGC 607
Db 644 TCTCACCCCTTCTCAGGCGGTGATGGGCTCTCTCATAGCGATTCAGATCTCTCTAAGC 703
Qy 608 AGCGGGTCAGTTCCTCGGTGAATACCTGGAAGCAAGAAATGCGCTATGGCTTCTCAT 667
Db 704 AGCGGGTCAGTTCCTCGGTGAATACCTGGAAGCAAGAAATGCGCTATGGCTTCTCAT 763
Qy 668 ATGACACCGCTGTTTGAATCAACCGGTCACTGAGAATGACATCCGCTGTGAGAGTCAA 727
Db 764 ATGACACCGCTGTTTGAATCAACCGGTCACTGAGAATGACATCCGCTGTGAGAGTCAA 823
Qy 728 TTATCAATGTTGACCTTGGCCCCCGAGAGCTAGACAGGCCATAGGCTGCTACAGAGC 787
Db 824 TCTACCAATGTTGACTTGGCCCCCGAGAGCTAGAGCTAGGCTATTAAGTCTGCTACAGAGC 883
Qy 788 GSCTCTATGTCGGGGTCCCAATGACTACTCCAAAGGCGAGAACTGCGGGCTATCGCGGT 847
Db 884 GSCTCTATGTCGGGGTCCCAATGACTACTCCAAAGGCGAGAACTGCGGGCTATCGCGGT 943
Qy 848 GCGCGCGAGCGGCTGCTGACGACTAGCTGCGGTAAATACCTCAGCTGCTTGAAGG 907
Db 944 GCGCGCGAGCGGCTGCTGACGACTAGCTGCGGTAAATACCTCAGCTGCTTGAAGG 1003
Qy 908 CGCTGACGCTGTCAGCTCCAGCTCCAGGACTGACAGATGCTCGTAATGGAGAGC 967
Db 1004 CCGTGGCGCTGTCGAGCTGCGAGCTCCAGGACTGACAGATGCTCGTAATGGAGAGC 1063
Qy 968 ACCTTGTGTTATCTGTGAAGCGGGGAAACCAAGAGGACCGGCAAGCTTACAGTCT 1027
Db 1064 ACCTTGTGTTATCTGTGAGAGCGGGGAAACCAAGAGGATGCGGCAACCTACAGTCT 1123
Qy 1028 TCAGGAGCTATGACTAGTACTCTGCCCCCCCCCTGGGAGCCGCCCAACCGGAATACG 1087
Db 1124 TCAGGAGCTATGACTAGTACTCTGCCCCCCCCCTGGGAGCTCTGCCCAACCGGAATACG 1183
Qy 1088 ACTTGGAGCTGATTAACATCGTGTCTTCCATGTGTCGTCGACAGATGCTCTGCA 1147
Db 1184 ACTTGGAGTGTAAACATCATGCTCTTCCATGTGTCGTCGACAGATGCTCTGCA 1243
Qy 1148 AAAGGTGTACTACTCAACCGTGACCCCAACCGTCCCGCTTGGCGGGCTGCGTGGAGA 1207
Db 1244 AAAGGTGTACTACTCAACCGTGACCCCAACCGTCCCGCTTGGCGGGCTGCGTGGAGA 1303
Qy 1208 CAGTAGGACACATCCAGTCACTCTGCTAGGCAACATCATGTATGCGGCCACTT 1267
Db 1304 CAGTAGACACATCCATCAACTCACTGCTAGGCAATATCATCATGTATGCGGCCACTT 1363
Qy 1268 TGTGGCAAGATGATTCGTATGACTCACTTCTCTCCATCTCTAGCCCGAGAGCAAC 1327
Db 1364 TATGGCAAGATGATTCGTATGACTCACTTCTCTCCATCTCTAGCCCGAGAGCAAC 1423
Qy 1328 TTGAAAAGCCCTGGATTTGTGAGATCTACAGGGGCTTGTACTCCATTTAGGCACTTGCAC 1387
Db 1424 TTGAAAAGCCCTGGATTTGTGAGATCTACAGGGGCTTGTACTCCATTTAGGCACTTGCAC 1483
Qy 1388 TACTCAGATCAATGAAGCACTCCATGCTTTAGGCACTTTTCACTCCATGTTACTCTC 1447
Db 1484 TACTCAGATCAATGAAGCACTCCATGCTTTAGGCACTTTTCACTCCATGTTACTCTC 1543
Qy 1448 CAGGTGAGATCAATAGGGTGTCTTATGCTTACAGGAAGCTTGGGGTACCAACCTTGGAG 1507
Db 1544 CAGGTGAAATCAATAGGGTGTCTTATGCTTACAGGAAGCTTGGGGTACCAACCTTGGAG 1603

Qy 1508 TCTGAGACATCGGGCCAGAGTGTCCGCGTAAAGTTACTGTCCAGGGGGAGGGCG 1567
Db 1604 TCTGAGACATCGGGCCAGAGTGTCCGCGTAAAGTTACTGTCCAGGGGGAGGGCG 1663
Qy 1568 CCATTTGTGGCAAGTACCTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCAA 1627
Db 1664 CCATTTGTGGCAAGTACCTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCAA 1723
Qy 1628 TCCGCGCTGGTCCGCGTGGACTTGTCCGCGTGGTTCGTTGCTGGCTACAGCGGGGAG 1687
Db 1724 TCCGCGCTGGTCCGCGTGGACTTGTCCGCGTGGTTCGTTGCTGGCTACAGCGGGGAG 1783
Qy 1688 ACATATATCAGAGCTGCTCGTCCGCGACCCCGCTGGTTCATGTTGCTTACTCTCTAC 1747
Db 1784 ACATATATCAGAGCTGCTCGTCCGCGACCCCGCTGGTTCATGTTGCTTACTCTCTAC 1843
Qy 1748 TCTCCGTAGGGGTAGGCACTTACTCTCCCAACCG 1784
Db 1844 TTCTGTAGGGGTAGGCACTTACTCTCCCAACCG 1880

RESULT 4

ADR82190

ID ADR82190 standard; DNA; 9414 BP.

XX ADR82190;

AC ADR82190;

XX DT 16-DEC-2004 (first entry)

XX Hepatitis C virus DNA.

DE Hepatitis C virus.

KW antilipemic; cardiant; vasotropic; antiarteriosclerotic; antidiabetic;

KW cyostatic; anticonvulsant; nootropic; muscula; anti-HIV;

KW RNA interference; iRNA; antisense technology; lipid metabolism;

KW cholesterol imbalance; dyslipidaemia hypercholesterolaemia;

KW coronary artery disease; CAD; coronary heart disease; CHD;

KW atherosclerosis; hepatic glucose production;

KW glucose-metabolism-related disorder; diabetes; cancer; breast cancer;

KW colon cancer; lung cancer; neurological disease; Huntington disease;

KW spinocerebellar ataxia; viral disease; AIDS; HCV; ds.

XX Hepatitis C virus.

OS Hepatitis C virus.

XX WO2004080406-A2.

XX 23-SEP-2004.

XX 08-MAR-2004; 2004WO-US0007070.

XX 07-MAR-2003; 2003US-0452682P.

XX 12-MAR-2003; 2003US-0454265P.

XX 13-MAR-2003; 2003US-0454562P.

XX 13-MAR-2003; 2003US-0455050P.

XX 14-APR-2003; 2003US-0462894P.

XX 17-APR-2003; 2003US-0463772P.

XX 25-APR-2003; 2003US-0465665P.

XX 25-APR-2003; 2003US-0465802P.

XX 09-MAY-2003; 2003US-0469612P.

XX 08-AUG-2003; 2003US-0493986P.

XX 11-AUG-2003; 2003US-0494597P.

XX 26-SEP-2003; 2003US-0506341P.

XX 09-OCT-2003; 2003US-0510246P.

XX 10-OCT-2003; 2003US-0510318P.

XX 07-NOV-2003; 2003US-0518453P.

XX (ALNY-) ALNYLAM PHARM.

XX Manoharan M, Bumcrot D;

XX WPI; 2004-677362/66.

XX Interference RNA agent useful for treating dyslipidemias, coronary artery

PT

PT disease, diabetes, cancer or neurological disease, comprises sense
sequence and antisense sequence which has specific modifications.

PS Example 5; SEQ ID NO 6889; 378pp; English.

XX The invention describes a RNA interference (iRNA) agent (I) comprising a
sense sequence and an antisense sequence, where the sense sequences have
one or more asymmetrical 2'-O alkyl modifications, the antisense
sequences have one or more asymmetrical phosphorothioate modifications
and the antisense sequence targets a human gene sequence. Also described
are: a pharmaceutical preparation comprising (I); reducing (MI) apob-100
levels or glucose-6-phosphatase levels in a subject; producing (I);
stabilising (I), involves selecting a sequence with activity and
introducing one or more asymmetrical modification in the sequence, where
the modification decreases nuclease sensitivity while not decreasing its
activity; a kit comprising (I) and instruction for its use; and a device
that can be dispense or administer a composition comprising (I). (I) is
useful for reducing apob-100 levels or glucose-6-phosphatase levels. (MI)
is useful for reducing apob-100 levels or glucose-6-phosphatase levels.
The subject is suffering from a disorder characterised by elevated or
otherwise unwanted expression of apob-100, elevated or otherwise unwanted
levels of cholesterol, and/or dysregulation of lipid metabolism. The
disorder is chosen from the HDL/LDL cholesterol imbalance,
dyslipidaemias, hypercholesterolaemia, statin-resistant
hypercholesterolemia, coronary artery disease (CAD), coronary heart
disease (CHD) and atherosclerosis. (I) is administered to a subject to
inhibit hepatic glucose production or for treating glucose-metabolism-
related disorder e.g. diabetes or type-2 diabetes. (I) is useful for
treating the diseases as mentioned above, cancer (e.g. breast, colon or
lung cancer), neurological disease (e.g., Huntington disease or
epinocerebellar ataxia) or viral disease (e.g., AIDS). This sequence
represents hepatitis C virus DNA.

XX Sequence 9414 BP; 1887 A; 2825 C; 2687 G; 2015 T; 0 U; 0 Other;

Query Match 89.2%; Score 1607.4; DB 13; Length 9414;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY	6	GCATGTCAATGTCCTATACATGACAGCGCGCCTGATCACACCGCTCGCGGAGGAAA	67
DB	7595	GCTGTTCCGATGCTTACATCGACAGCGCGCTGATCACGCCATGCGCTCGCGAGGAAA	7654
QY	68	GCAAGTCGCCATCAACCGCTGAGCAACTCTTCCTGCGTCACCATAACTGCTCTATT	127
DB	7655	GCAAGTTGCCATCAACCGCTTGAGCAATCTTCTGCTACGTCACCAACAATGCTCTATG	7714
QY	128	CCACAACATCCCGAGTGAAGCCTGCGGCACGAAGAGTTCACCTTCACAGACTGCAAG	187
DB	7715	CTACAACATCCCGCAGCGCAGGCTCGGCACGAAGAAGTTCACCTTTGACAGACTGCAAG	7774
QY	188	TCCTGGACCATATTACCGGACGCTCTCAAGGAGATGAAGCGGAAGCGCTCCAGTGA	247
DB	7775	TCTCGACACCACTACCGGACGCTCTTAAGGAGATGAAGCGGAAGCGCTCCACGTGA	7834
QY	248	AGCTAAACTGCTATCTGTAGAAGAGCATGCAAGCTCAACGCCCGCATTCGGCCAAAT	307
DB	7835	AGCTAAACTTCTATCTGTAGAAGAGCGCTGCAAACTGACGCCCGCCACATTCGGCCNAAT	7894
QY	308	CCAAATTTGGCTATGGGCAAGAGCGTCCGAGGCTATCCAGCAGGCGGTAAACCAACA	367
DB	7895	CCAAATTTGGCTACGGGCGAAGGACGTCGGACCCCTATCCAGCAGGCGGTAAACCAACA	7954
QY	368	TCGCTCCGTTGGAGAGACTTGTCTGGAGCACACTGACACCAATTCAGACCAATCA	427
DB	7955	TCGCTCCGTTGGAGAGACCTGCTGGAGACACACTGAAACACCAATTAAGCACTACATCA	8014
QY	428	TGGCAAAAATGAGTTTCTGGTCCAAACAGAGAAAGCGCGCAAAACAGCTCGCC	487
DB	8015	TGGCAAAAATGAGTTTCTGTTCCAAACAGAGAAAGCGCGCAAGCCAGCTCGCC	8074
QY	488	TCATCGTATTTCCAGACCTGGGAGTTCGTGTATGCGAGAAGATGGCCCTTTACGACGTGG	547

DB	8075	TTATCGTGTCCAGATCTGGGAGTTCGTGTATGCGAGAGATGCCCTTTATGACGTGG	8134
QY	548	TTTCCACTCTTCTCAGGCGGTGATGGGCTCCTCATACGGATTCCAATATCTCTCTAAGC	607
DB	8135	TCTCCACCTTCTCAGGCGGTGATGGGCTCCTCATACGGATTCCTCCTCTAAGC	8194
QY	608	AGCGGTCGAGTCTCTGGTGGAATACCTGGAAGCAAGAAATGCCCTATGGCTTCTCAT	567
DB	8195	AGCGGTCGAGTCTCTGGTGGAATACCTGGAAGCAAGAAATGCCCTATGGCTTCTCAT	8254
QY	668	ATGACACCCGCTGTTTGTACTCAACGGTCACTGAGNATGACATCCGTGTTGAGGAGTCAA	727
DB	8255	ATGACACCCGCTGTTTGTACTCAACGGTCACTGAGNATGACATCCGTGTTGAGGAGTCAA	8314
QY	728	TTTACCNAATGCTGACTTGGCCCCGGAAGCTAGACAGGCCATTAAGTTCGCTCACAGAGC	787
DB	8315	TTTACCNAATGCTGACTTGGCCCCGGAAGCCAACTGGCCATTAAGTTCGCTCACAGAGC	8374
QY	788	GGCTCTATGTCGGGGGCTCCCATGACTAACTCCAAAGGGCAGAACTCGGCTATFCGCCGT	847
DB	8375	GGCTCTATGTCGGGGGCTCCCTGACTTAATTCAAAGGGCAGAACTCGGCTATFCGCCGT	8434
QY	848	GCCGCGCAGCGCGTCTGACGACTAGCTGCGGTAAATACCTCACATGTTACCTGAAAG	8494
DB	8435	GCCGCGCAGCGCGTCTGACGACTAGCTGCGGTAAATACCTCACATGTTACCTGAAAG	8494
QY	908	CCGCTGACGCTCTCGAGCTGCCAAGCTCCAGACATGACGATGCTCGTGAATGGAGAGC	967
DB	8495	CCACTCGGCTCTGCGAGCTGCGAAGCTCGGAGCTGCGAGCTCTGTTGAACGAGAGC	8554
QY	968	ACCTTGCTGTTATCTGTAAAGCGCGGAAACCCAGAGGAGCGGGAAGCTTACGAGTCT	1027
DB	8555	ACCTTGCTGTTATCTGTAAAGCGCGGGAACCCAGAGGATGCGGCGAGCTTACGAGTCT	8614
QY	1028	TCACGGAGGCTATGACTAGGTACTCTGCCCCCTCTGGGGACCCGCCCAACCGGAATACG	1087
DB	8615	TCACGGAGGCTATGACTAGGTACTCTGCCCCCTCTGGGGACCCGCCCTCAACCGGAATACG	8674
QY	1088	ACTTGGAGCTGATAACATCGTGTCTCCAAATGCTGCGTGCACACGATGCATCTGGCA	1147
DB	8675	ACTTGGAGTGTATAACATCATGTTCTCCAAATGCTGCGTGCACACGATGCATCTGGTA	8734
QY	1148	AAAGGTTGTTACTCTCACCCGTGACCCACCTGCGGCGGCTGCGTGGGAGA	1207
DB	8735	AAAGGTTGTTACTCTCACCCGTGACCCACCTTACCAACCCCTTGCACGGCTGCGTGGGAGA	8794
QY	1208	CAGCTAGGCACTCCAGTCAAATCTCTGGCTAGGCAACATCATCATGATGTCGCCACCT	1267
DB	8795	CAGCTAGACACATCCAGTCAAATCTCTGGCTAGGCAACATCATCATGATGTCGCCACCT	8854
QY	1268	TGTGGGCAAGATGATTTGATGATCACTTTCTTCTCCATCTCTTAGGCCAGGAGCAAC	1327
DB	8855	TATGGGCAAGATGATTTGATGATCACTTTCTTCTCCATCTCTTAGCTCAGGAGCAAC	8914
QY	1328	TTGAAAAGCCCTGGATTTGATGATCACTTTCTTCTCCATCTCTTAGCTCAGCTTACCT	1387
DB	8915	TTGAAAAGCCCTTAGATTTGATGATCACTTTCTTCTCCATCTCTTAGCTCAGCTTACCT	8974
QY	1388	TACCTCAGATCATTTGAACGACTCCATGCTTTAGCGCAATTTTCACTCCATAGTTACTCTC	1447
DB	8975	TACCTCAGATCATTTGACGACTCCATGCTTTAGCGCAATTTTCACTCCATAGTTACTCTC	9034
QY	1448	CAGGTGAGATCAATAGGCTGCTTTCATGCTCAGAAAGCTTGGGGTACCACTTGGGAG	1507
DB	9035	CAGCGAGATCAATAGGCTGCTTTCATGCTCAGAAATCTGGGGTACCACTTGGGAG	9094
QY	1508	TCTGGAGACATCGGCGCAGAGTGTCCGCGCTTAAGTTACTGTCCAGGGGGAGGGCGG	1567
DB	9095	CCTGGAGACATCGGCGCAGAGTGTCCGCGCTTAAGTTACTGTCCAGGGGGAGGGCGG	9154
QY	1568	CCATTTGTCGCAAGTACTCTTCAACTGGGCGTAAGGACCAAGCTTAAACTCTACTCCAA	1627
DB	9155	CCACTTGTGGCAAGTACTCTTCAACTGGGCGGTAGGACCAAGCTTAAACTCTACTCCAA	9214

Qy	1628	TTCCGGCTGCGTCCCGGCTGGAATTGTCGGCTGGTTTCGTTGGCTACAGCGGGGAG	1688
Db	9215	TCCAGCGCGGTCCCGGTTGGACTTGTTCGGCTGGTTTCGTTGGCTTACAGCGGGGAG	9274
Qy	1688	ACATATATACAGCGCTGCTCTCGTGCCCGACCCCGCTGGTTTCATGTTGTGCTACTCTCTAC	1747
Db	9275	ACATATATACAGCGCTGCTCTCGTGCCCGACCCCGCTGGTTTCATGTTGTGCTACTCTCTAC	9334
Qy	1748	TCTCCGTAGGGTAGGCATCTATCTACTCTCCCAACCG	1784
Db	9335	TTTCCGTGGGGTAGGCATCTACTCTCTCCCAACCG	9371
RESULT 5			
AAQ33282			
ID	AAQ33282	standard; cDNA; 9472 BP.	
AC	AAQ33282;		
XX			
DT	27-AUG-2003	(revised)	
DT	25-MAR-2003	(revised)	
DT	19-MAY-1993	(first entry)	
XX			
DE		Korean hepatitis C virus full cDNA sequence LBC1.	
XX			
KW		KHCV-LBC1; diagnosis; vaccine; ds.	
XX			
OS		Hepatitis C virus.	
XX			
Key		Location/Qualifiers	
FT	CDS	343..9375	
FT		/*tag= a	
FT		/note= "KHCV polypeptide"	
XX			
PN	EPS21318-A2.		
XX			
PD	07-JAN-1993.		
XX			
PF	10-JUN-1992;	92EP-00109753.	
XX			
PR	10-JUN-1991;	91KR-00009510.	
PR	06-AUG-1991;	91KR-00013601.	
XX			
XX	(LUCK-) LUCKY LTD.		
PA			
XX			
PI	Cho JM, Lee YB, Park YW, Lim KJ, Choi DY, So HS, Kim CH;		
PI	Kim ST, Yang JY;		
XX			
DR	WPI; 1993-001883/01.		
DR	P-PSDB; AAR30616.		
XX			
PT		DNA and polypeptide(s) from a new type of hepatitis C virus (KHCV) - for	
PT		diagnosing and vaccinating against KHCV infections.	
XX			
PS	Claim 7; Fig 2; 119pp; English.		
XX			
CC		This is the full cDNA sequence of Korean hepatitis C virus (KHCV) cDNA,	
CC		KHCV-LBC1. Portions of the cDNA may be used as probes or primers to	
CC		diagnose the presence of the virus in putative samples. Polypeptides	
CC		derived from the cDNA may be used in a specific and accurate method for	
CC		detecting KHCV antibodies in the serum of hepatitis C patients.	
CC		Antibodies directed against these polypeptides are useful for the	
CC		purification of KHCV antigens and for the development of an improved	
CC		diagnostic to detect KHCV antigens in a sample. The polypeptides may also	
CC		be used in a vaccine for treatment and prevention of KHCV infection at a	
CC		dosage of 5-200 ug/peptide. (Updated on 25-MAR-2003 to correct PN field.)	
XX		(Updated on 27-AUG-2003 to correct OS field.)	
SQ		Sequence 9472 BP; 1887 A; 2826 C; 2686 G; 2073 T; 0 U; 0 Other;	


```
QY 1088 ACTTGGAGCTGATAACATCGTGTGTTCTCTCCAAATGTTGGTCGCACACGATGCAATCTGGCA 1147
Db ACTTGGAGTTGATAACATCATGTTCTCTCCAAATGTTGGTCGCACACGATGCAATCTGGTA 8734
QY 1148 AAAGGGTGTACTACTCACTACCCGTCGACCCACCGTCCCTTGGCGGGTGGCGGAGAGA 1207
Db AAAGGGTGTACTACTCACTACCCGTCGACCCACCGTCCCTTGGCGGGTGGCGGAGAGA 8794
QY 1208 CAGCTAGGCACACTCCAGTCAACTCTGGCTAGGCAACATCATCATGTATGCGCCCACTT 1267
Db CAGCTAGGCACACTCCAGTCAACTCTGGCTAGGCAACATCATCATGTATGCGCCCACTT 8854
QY 1268 TGTGGGCAAGGATGATCTGATGACTCACTTCTCTCCATCTCTTAGCCCCAGGAGCAAC 1327
Db TATGGGCAAGGATGATCTGATGACTCACTTCTCTCCATCTCTTAGCTCAGGAGCAAC 8914
QY 1328 TTGAAAAAGCCCTGGATTTGTCAGATCTACGGGGCTTTGCTCAATTAAGCACTTGCAAC 1387
Db TTGAAAAAGCCCTAGATTTGTCAGATCTACGGGGCTTTGCTCAATTAAGCACTTGATC 8974
QY 1388 TACCTCAGATCATTGAACGACTCCATGCTCTAGGCAATTTTCACTCCATAGTTACTCTC 1447
Db TACCTCAGATCATTGAGCGACTCCATGCTCTTAGCGCAATTTTCACTCCATAGTTACTCTC 9034
QY 1448 CAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTTGGGGTACCACCTTGGGAG 1507
Db CAGGCGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTTGGGGTACCACCTTGGGAG 9094
QY 1508 TCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGCGCG 1567
Db CTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGCGCG 9154
QY 1568 CAATTTGTGGCAAGTACCTCTTCAACTGGGCAAGTAAAGCAACAACTTAACTCACTCAA 1627
Db CCACCTTGTGGCAAGTACCTCTTCAACTGGGCGGTGAGGACCAAGCTCAAACTCACTCAA 9214
QY 1628 TTCGGCTCGTCCGGCTGGACTGTCCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1687
Db TCCAGCGCGTCCGGCTGGACTGTCCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 9274
QY 1688 ACATATATACAGGCTGTCTCGTCCGACCCGCTGGTTCATGTTGTCCTACTCTCTAC 1747
Db ACATATATACAGGCTGTCTCGTCCGACCCGCTGGTTCATGTTGTCCTACTCTCTAC 9334
QY 1748 TCTCGTAGGGGTAGGCATCTATCTACTCTCCCAACCG 1784
Db TTTCCGTGGGGTAGGCATCTACTCTCTCTCCCAACCG 9371
```

RESULT 6

ACA61114

ID ACA61114 standard; cDNA; 5955 BP.

XX ACA61114;

XX 09-JUL-2003 (first entry)

XX HCV cDNA encoding Met-NS3-NS4A-NS4B-NS5A-NS5B (active).

XX HCV; ss; gene; non-structural protein; NS3; NS4A; NS4B; NS5A; NS5B;
XX adenoviral vector; HCV infection; vaccine; gene therapy; protease.
XX Hepatitis C virus.

XX Key Location/Qualifiers

FH 1..5955

FT /*tag= a

FT /product= "Met-NS3-NS4A-NS4B-NS5A-NS5B"

FT /partial

FT /note= "No stop codon shown"

XX

XX WO2003031588-A2.

XX 17-APR-2003 .
XX 10-OCT-2002; 2002WO-US032512.
XX 11-OCT-2001; 2001US-0328655P.
XX 13-MAR-2002; 2002US-0363774P.
XX (MERI) MERCK & CO INC.
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX Emni EA, Kaslow DC, Bett AJ, Shiver JW, Nicosia A, Lahm A;
XX Luzzago A, Cortese R, Colloca S;
XX WPI; 2003-381708/36.
XX P-PSDB; ABU09575.
XX New nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A-NS5B polypeptide,
XX useful as a component of an adenovector or DNA plasmid vaccine for
XX preventing or treating hepatitis C virus.
XX Example 1; Fig 5; 231pp; English.
XX The invention relates to a nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A-
XX -NS5B (NS stands for non-structural protein from HCV) polypeptide
XX appearing as ABU09574. The encoded polypeptide has sufficient protease
XX activity to process itself to produce an NS5B protein that is
XX enzymatically inactive. Also included are a cultured recombinant cell
XX comprising the novel nucleic acid, making an adenovector (comprising: (a)
XX producing an adenovirus genome plasmid comprising a gene expression
XX cassette by homologous recombination between the novel nucleic acid and a
XX nucleic acid comprising a first adenovirus region from base pair 1-450
XX corresponding to either Ad5 or Ad6, a second adenovirus region from base
XX pair 3511-5548 corresponding to Ad5 or from base pair 3508-5541
XX corresponding to Ad6, joined to the first region, a third adenovirus
XX region from base pair 5549-28133 corresponding to Ad5 or from base pair
XX 5542-28156 corresponding to Ad6, joined to the second region, a fourth
XX adenovirus region from base pair 30818-33966 corresponding to Ad5 or from
XX base pair 30789-33784 corresponding to Ad6, joined to the third region,
XX and a fifth adenovirus region from base pair 33967-35935 corresponding to
XX Ad5 or from base pair 33785-35759 corresponding to Ad6, joined to the
XX fourth region; and (b) rescuing the recombinant adenovirus from the
XX recombinant adenovirus plasmid), an adenoviral vector that is produced by
XX method above appearing as ACA61113 which has a humanised version of the
XX polynucleotide of the invention and encodes the HCV inactivated
XX polypeptide, a recombinant nucleic acid comprising one or more Ad6
XX regions and a region not present in Ad6, where at least one Ad6 region is
XX selected from E1A, E1B, E2B, E2A, E4, L1, L2, L4 and L5. The nucleic acid
XX is useful as a component of an adenoviral vector or a DNA plasmid vaccine
XX providing a broad range of antigens for generating an HCV-specific cell
XX mediated immune response. The nucleic acid may also be used in treating
XX patients infected with HCV. The present sequence is the non-humanised
XX cDNA encoding the wild-type HCV polypeptide with an active NS5B protein
XX
XX Sequence 5955 BP; 1251 A; 1798 C; 1678 G; 1228 T; 0 U; 0 Other;

Query Match

Best Local Similarity 88.6%; Score 1597.8; DB 8; Length 5955;

Matches 1665; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 8 GCATGTCAATGTCTTATACATGACAGCGGCGCTGATCACACCGTGGCTGGCGGAGGMAA 67
Db 4178 GCTGCTCAATGTCTTACATGACAGCGGCGCTGATCACCGCATGGCTGGCGGAGAAA 4237

Qy 68 GCAAGCTGCCCATCAACGCGCTGAGCAACTCTTGTGCTGCCTACCACTAACTGCTCTATT 127
Db 4238 GCAAGCTGCCCATCAACGCGCTGAGCAACTCTTGTGCTGCCTACCACTAACTGCTCTATT 4297

Qy 128 CCACACATCCCCAGTGCAGGCTCGGCGAGAGAAGGTTCACCTTTTGACAGACTGCAAG 187
Db 4298 CCACACATCTCGCAGCGCAGGCTCGGCGAGAGAAGGTTCACCTTTTGACAGACTGCAAG 4357

Qy 188 TCCTGGACGATCATTTACCGGGACGTGCTCAAGAGATGAAGCGGAGGCGTCCACAGTGA 247

Db	4358	TCCTGGACGACACCTACCGGACCGTGCTCAAGAGAGNTGAAGGCGAAGGCGGTCCACAGTTA	4417
Qy	248	AGGCTAAACTGCTATCTGTAGAAGAGCATCAAGCTGACGCCGCCGCCCATTTGCGCCAAAT	307
Db	4418	AGGCTAAACTCCTATCCGTAGAGAGAGCGTGCAAGCTGACGCCGCCCCACATTCGCGCAAAAT	4477
Qy	308	CCAAATTTGGCTATGGGCAAGAGCAGCTCGGAGCCCTATCCAGCAGGCGCGTTAACCAACA	367
Db	4478	CCAAAGTTGGCTATGGGCGCAAGAGCAGCTCGGAAACCTATCCAGCAAGCGCGTTAACCAACA	4537
Qy	368	TCCGCTCGGTGTGAAGGACCTTGCTGTGAGGACACTGTACACACCAATTCAGACCAACCATCA	427
Db	4538	TCCACTCCGCTGTGAAGGACCTTGCTGTGAAGACACTGTGTGACACCAATTTGACACCAACATCA	4597
Qy	428	TGGCAAAAATAGAGTTTTCGCGTCCAAACGAGAGAAAGGAGCGCGCAAAACCAAGCTCGCC	487
Db	4598	TGGCAAAAATAGAGTTTTCGCGTCCAAACGAGAGAAAGGAGCGCGTAAGCCAGCCCGCC	4657
Qy	488	TCATCGTATCCAGACCTGGAGTTCGTGTATCGGAGAGATGGCCCTTTACGACGTGG	547
Db	4658	TTATCGTATTCAGATCTGGAGTCCGTTATCGGAGAGATGGCCCTCTATGATGTGG	4717
Qy	548	TTTCCACTTTCCTCAGGCCGTGATGGCTCCTCATACGGATTCCCAATACTCTCTCCTAAGC	607
Db	4718	TCGCCACCTTCCTCAGGTCTGATGGCTCCTCATACGGATTCAGTACTCTCTGGGC	4777
Qy	608	AGCGGTGCGATTCTCTGGTGAATACCTGGAAAGCAAAAGAAATGCCCTATGGGCTTCTCAT	667
Db	4778	AGCGAGTCGAGTTCTCTGGTGAATACCTGGAAATCAAAGAAAACCCCATGGGCTTTTCAT	4837
Qy	668	ATGACACCGCGTGTTTTGACTCAACCGTCACTGAGAATGACATCCGTGTTGAGGAGTCAA	727
Db	4838	ATGACACTCGCTGTTTCGACTCAACGCTCACCGAGAACGACATCCGTGTTGAGGAGTCAA	4897
Qy	728	TTTACCATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCCATTAAGTCTCCTCAGAGAC	787
Db	4898	TTTACCAATGTTGTGACTTGGCCCCCGAAGCCAGACAGGCCATAAAATCGCTCAGAGAC	4957
Qy	788	GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGACTCGGGCTATCGCGGT	847
Db	4958	GGCTTTTATATCGGGGTCTCTGACTAAATTCAAAAGGGCGAGACTCGGTATTCGCGGT	5017
Qy	848	GCCGCGGAGCGCGGTGCTGACGACTAGCTGCGGTATACCCCTCACATGCTACTTGAAGG	907
Db	5018	GCCGCGGAGCGCGGTGCTGACGACTAGCTGCGGTATACCCCTCACATGTTACTTGAAGG	5077
Qy	908	CCGCTGAGGCTGTGAGCTGCCAAGCTCCAGACTGACAGATGCTCGTGAATGGAGACG	967
Db	5078	CCTCTGAGGCTGTGAGCTGCGAAGCTCCAGACTGACAGATGCTCGTGAACGGAGACG	5137
Qy	968	ACCTTGTGTTATCTGTGAAGCGCGGAAACCCAGAGGACGCGGCAAGCCTACGAGTCT	1027
Db	5138	ACCTTGTGTTATCTGTGAAGCGCGGAAACCCAGAGGACGCGGCGAGCCTACGAGTCT	5197
Qy	1028	TCAGGAGGCTATGACTAGTACTCTGCCCCCTGGGACCGCGGCCAACCGGAATACG	1087
Db	5198	TCAGGAGGCTATGACTAGTACTCTGCCCCCTGGGACCGCGGCCAACCGGAATACG	5257
Qy	1088	ACTTGGAGCTGATAAATCTGTGTTCTCCAAATGTGCGTCCGACACGATGATCTGGCA	1147
Db	5258	ACTTGGAGCTGATAAATCTGTTCTCCAAATGTGCGTCCGACACGATGATCTGGCA	5317
Qy	1148	AAAGGGTGTACTACTCTCAACCGGTGACCCACCGTCCCCCTTGGCGGGCTGGCGGAGA	1207
Db	5318	AAAGGGTGTACTACTCTCAACCGGTGATCCCAACCGTCCCCCTTGGCGGGCTGGCGGAGA	5377
Qy	1208	CAGCTAGGCACTCTCCAGTCAACTCTCTGGCTAGGCAACATCATCATGTATGCGCCCATTT	1267
Db	5378	CAGCTAGCACACTCTCCAGTTAACTCTCTGGCTAGGCAACATTTATCATGTATGCGCCCATTT	5437
Qy	1268	TGTTGGCAAGGATGATTTCTGATGACTCATCTTCTTCTCATCTTCTAGGCCAGGAGCAAC	1327

Db	5438	TGTGGCAAGATGATTCTGATCACTCAGTCTTCTCTCCATCTCTTAGCACAGGAGCAAC	5497
Qy	1328	TTGAAAGAGCCCTGGATTTGTCAGATCTACGGGCTTGTACTCTCATTTAGCCACTTGACC	1387
Db	5498	TTGAAAGAGCCCTGGAGCTGCAGATCTACGGGCTTGTACTCTCATTTAGCCACTTGACC	5557
Qy	1388	TACCTCAGATCAATTGAACGACTCCATGGTCTTAGCCGCAATTTTCACTCCATAGTTACTCTC	1447
Db	5558	TACCTCAGATCAATTGAACGACTCCATGGCTTAGCGCATTTTCACTCCATAGTTACTCTC	5617
Qy	1448	CAGGTGAGATCAATAGGTGGCTTCATGCCTCAGGAAGCTTTGGGGTACCAACCTTGGCAG	1507
Db	5618	CAGGTGAGATCAATAGGTGGCTTCATGCCTCAGGAAGCTTTGGGGTACCAACCTTGGCAG	5677
Qy	1508	TCTGGAGACATCGGGCCAGAGTGTCCGGCTTAGTACTGTCCAGGGGGGAGGGCCG	1567
Db	5678	TCTGGAGACATCGGGCCAGAGGCTCCGGCTAGGCTACTGTCCAGGGGGGAGGGCCG	5737
Qy	1568	CCATTTGTGCAAGTAGTACCTCTTCAACTGGGCGTAGGACCAAGCTTAAACTCAGTCCAA	1627
Db	5738	CCACTTTGTGCAAGTAGTACCTCTTCAACTGGGCGTAGGACCAAGCTTAAACTCAGTCCAA	5797
Qy	1628	TTCCGGCTGGCTCCCGCTGGACTTGTCCGGCTGGTTGCTTGTGGCTACAGCGGGGAG	1687
Db	5798	TCCGGCTGGCTCCCGCTGGACTTGTTCGGCTGGTTGCTTGTGGCTACAGCGGGGAG	5857
Qy	1688	ACATATATCACAGCCTGTCTCGTGCCCGACCCCGCTGGTTTCATGTTGTGCTACTCTTAC	1747
Db	5858	ACATATATCACAGCCTGTCTCGTGCCCGACCCCGCTGGTTTCATGTTGTGCTACTCTTAC	5917
Qy	1748	TCTCCGTAGGGGTAGGATCTATCTACTCCCCAACG	1784
Db	5918	TTTCTGTAGGGGTAGGATCTACTCTCTCCCCAACG	5954
RESULT 7			
ADR82189			
ID	ADR82189	standard; DNA; 9587 BP.	
XX	AC	ADR82189;	
XX	DT		
XX	DT	16-DEC-2004 (first entry)	
DE	DE	Hepatitis C virus type 1b polyprotein DNA.	
XX	KW	antilipemic; cardiant; vasotropic; antiarteriosclerotic; antidiabetic;	
KW	KW	cytostatic; anticonvulsant; nootropic; muscular; anti-HIV;	
KW	KW	RNA interference; iRNA; antisense technology; lipid metabolism;	
KW	KW	cholesterol imbalance; dyslipidaemia hypercholesterolaemia;	
KW	KW	coronary artery disease; CAD; coronary heart disease; CHD;	
KW	KW	atherosclerosis; hepatic glucose production;	
KW	KW	glucose-metabolism-related disorder; diabetes; cancer; breast cancer;	
KW	KW	colon cancer; lung cancer; neurological disease; Huntington disease;	
KW	KW	spinocerebellar ataxia; viral disease; AIDS; HCV; polyprotein; gene; ds.	
XX	OS	Hepatitis C virus.	
XX	PN		
XX	PN	WO2004080406-A2.	
XX	XX		
PD	PD	23-SEP-2004.	
XX	XX		
XX	XX	08-MAR-2004; 2004WO-US0007070.	
PR	PR	07-MAR-2003; 2003US-0452682P.	
PR	PR	12-MAR-2003; 2003US-0454265P.	
PR	PR	13-MAR-2003; 2003US-0454962P.	
PR	PR	13-MAR-2003; 2003US-0455050P.	
PR	PR	14-APR-2003; 2003US-0462894P.	
PR	PR	17-APR-2003; 2003US-0463772P.	
PR	PR	25-APR-2003; 2003US-0465665P.	
PR	PR	25-APR-2003; 2003US-0465802P.	
PR	PR	09-MAY-2003; 2003US-0469612P.	
PR	PR	08-AUG-2003; 2003US-0493986P.	

11-AUG-2003; 2003US-0494597P.
26-SEP-2003; 2003US-0506341P.
09-OCT-2003; 2003US-0510246P.
10-OCT-2003; 2003US-0510318P.
07-NOV-2003; 2003US-0518453P.
(ALNY-) ALNYLAM PHARM.
Manoharan M, Bumcrot D;
WPI; 2004-677362/66.
Interference RNA agent useful for treating dyslipidemias, coronary artery disease, diabetes, cancer or neurological disease, comprises sense sequence and antisense sequence which has specific modifications.
Example 5; SEQ ID NO 6688; 378pp; English.
The invention describes a RNA interference (iRNA) agent (I) comprising a sense sequence and an antisense sequence, where the sense sequences have one or more asymmetrical 2'-O alkyl modifications, the antisense sequences have one or more asymmetrical phosphorothioate modifications and the antisense sequence targets a human gene sequence. Also described are: a pharmaceutical preparation comprising (I); reducing (M1) apoB-100 levels or glucose-6-phosphatase levels in a subject; producing (I); stabilising (I), involves selecting a sequence with activity and introducing one or more asymmetrical modification in the sequence, where the modification decreases nuclease sensitivity while not decreasing its activity; a kit comprising (I) and instructions for its use; and a device that can be dispense or administer a composition comprising (I). (I) is useful for reducing apoB-100 levels or glucose-6-phosphatase levels. (M1) is useful for reducing apoB-100 levels or glucose-6-phosphatase levels. The subject is suffering from a disorder characterised by elevated or otherwise unwanted expression of apoB-100, elevated or otherwise unwanted levels of cholesterol, and/or dysregulation of lipid metabolism. The disorder is chosen from the HDL/LDL cholesterol imbalance, dyslipidaemias, hypercholesterolaemia, statin-resistant hypercholesterolaemia, coronary artery disease (CAD), coronary heart disease (CHD) and atherosclerosis. (I) is administered to a subject to inhibit hepatic glucose production or for treating glucose-metabolism related disorder e.g. diabetes or type-2 diabetes. (I) is useful for treating the diseases as mentioned above, cancer (e.g. breast, colon or lung cancer), neurological disease (e.g., Huntington disease or spinocerebellar ataxia) or viral disease (e.g., AIDS). This sequence represents hepatitis C virus type 1b polyprotein DNA.
Sequence 9587 BP; 1321 A; 2865 C; 2714 G; 2087 T; 0 U; 0 Other;

7894 CCAAGTTTGGCTATGGGGCAAGAGAGTCCGGAACTATCCAGCAAGGCGGTAAACACA 7955

368 TCCGGTCCGGTGTGAAGGACTTGTCTGGAGGACACTGCACACACCAATTTCAGACCAACATCA 427

7954 TCCACTCCGTGTGAAGGACTTGTCTGAAGACACTGTGACACCAATTTGACACCAACATCA 8013

428 TGGCAAAAATGAGGTTTTCTGCGTCCAAACGACGAAGAGGCGGCAGAACCAAGCTCGCC 487

8014 TGGCAAAAATGAGGTTTTCTGTGTCCAAACGACGAAGAGGCGGCAGAACCAAGCTCGCC 8073

488 TCATCGTATTTCCACAGACTCGGAGTTCTGTGTATGCGAGAGATGSCCTTTTACGAGCTGG 547

8074 TTATCGTATTTCCAGATCTGGGAGTCCGTGTATGCGAGAGATGSCCTCTATGATGTGG 8133

548 TTTCACACTTCTCTCAGCGCGTGATGGGCTCCTCATACGGAATCCAAATCTCTCTAAGC 607

8134 TCTCCACCTTCTCTCAGGTCGTGATGGGCTCCTCATACGGAATCCAGTACTCTCTCGGC 8193

608 AGCGGTCGAGTTCTCTGCTGAATACCTGGAAGCAAAAGAAATTCGCCCTATGGGCTTCTCAT 667

8194 AGCGAGTCGAGTTCTCTGCTGAATACCTGGAATTCAAAGAGAAACCCCATGGGCTTTTCAT 8253

668 ATGACACCGGCTGTTTGACTCAACCGTCACCTGAGNAATGACATCCGTGTTTGAGGAGTCAA 727

8254 ATGACACCTGCTGTTTGCATCAACCGTCAACGAGAAACGACATCCGTGTTGAGGAGTCAA 8313

728 TTATACCAATTTGTGACTTTGGCCCCCGAAGCTAGACAGGCACTAAGGTCGCTCACAGAGC 787

8314 TTATACCAATTTGTGACTTTGGCCCCCGAAGCTAGACAGGCACTAAGTTCGCTCACAGAGC 8373

788 GGCTCTATGTCCGGGGTCCATGACTAACTTCCAAAGGCGAGAATCTGGGCTATCGCGGT 847

8374 GGCTTTATATCGGGGTCTCTGACTAACTTCCAAAGGCGAGAATCTGGGCTATCGCGGT 8433

848 GCGCGCGAGCGGCGTGTCTGACGACTAGCTGCGGTTAATCCTCAGTGTACTTTGAGG 907

8434 GCGCGCGAGCGGCGTGTCTGACGACTAGCTGCGGTTAATCCTCAGTGTACTTTGAGG 8493

908 CCGCTCAGCCCTGTCCAGCTGCCAAGCTCCAGAGCTCCAGAGTGTCTGTGAATGGAGACG 967

8494 CCTCTCAGCCCTGTCCAGCTGCCAAGCTCCAGAGCTCCAGAGTGTCTGTGAATGGAGACG 8553

968 ACCTTGTCTTATCTGTGAAGCGCGGAAACCCAGAGGACGCGCAAGCTTACGAGTCT 1027

8554 ACCTTGTCTTATCTGTGAAGCGCGGAAACCCAGAGGACGCGCGAGCTTACGAGTCT 8613

1028 TCACGAGGCTAATGACTAGGTACTCTGCGCCCCCTTGGGACCGGCCCCAACCGGAATACG 1087

8614 TCACGAGGCTAATGACTAGGTACTCTGCGCCCCCTTGGGACCGGCCCCAACCGGAATACG 8673

1088 ACTTGAGCTGATACATCGTTCTTCCAATGTGCGGTGCGACACGATGCATCTGCGCA 1147

8674 ACTTGAGCTGATACATCGTTCTTCCAATGTGCGGTGCGGCGGATGCATCTGCGCA 8733

1148 AAAGGCTGTACTACCTCAACCGCTGACCCACCGCTCCCCCTTGGCGGGCTCGTGGGAGA 1207

8734 AAAGGCTGTACTACCTCAACCGCTGATCCACACACCCCCCTTGGCAGCGGCTGGTGGGAAA 8793

1208 CAGCTAGGCACATCCAGTCAACTCTTGGCTAGGCAACATCATATGATGATGGGCCCACTT 1267

8794 CAGCTAGACACATCCAGTTAACTCTTGGCTAGGCAACATCATATGATGATGGGCCCACTT 8853

1268 TGTGGCAAGGATGATCTGATGACTCACTTCTTCTCCATCTCTTACGCCAGGAGCAAC 1327

8854 TGTGGCAAGGATGATCTGATGACTCACTTCTTCTCCATCTCTTACGCCAGGAGCAAC 8913

1328 TTGAAAAGCCCTCGAATTTGACAGTCTTACGCGGCTTGTGTACTCCATTTGAGCCACTTGACC 1387

8914 TTGAAAAGCCCTCGAATTTGACAGTCTTACGCGGCTTGTGTACTCCATTTGAGCCACTTGACC 8973

1388 TACCTCAGATCATTTGAACGACTCCATGGTCTTAGGCGATTTTTCATCTCCATGATTACTCTC 1447

908	CCGCTGCAGCCTGTGAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAATGGAGACG	967
Db		
	6878	CCACTGCAGCCTGTGAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAACGGAGACG
Qy		
	968	ACCTTGTCGTTATCTGTGAAAGCGCGGAAACCAAGAGACGGCGCAAGCCTACGAGTCT
Db		
	6938	ACCTTGTCGTTATCTGTGAAGCGGGAACCCAGGAGGATGCGCGGCCCTACGAGCCT
Qy		
	1028	TCACGAGGCTATGACTAGTACTCTGCCCCCCTGGGACCGCCCCCAACCGGAATACG
Db		
	6998	TCACGAGGCTATGACTAGTATGTTCCGCCGCCCGGGATCCGCCCAACCAAGATACG
Qy		
	1088	ACTTGGAGCTGATAACATCGTGTTCCTCCAAATGFTGCGTGGCACACGATGCATCTGGCA
Db		
	7058	ACCTGGAGCTGATAACATCATGTTCTCTCCAAATGTGTCGCGCAGATGCATCTGGCA
Qy		
	1148	AAAGGTTGTACTACCTCACCCGTGACCCACCCGTCCTTGGCGGGGTGGCTGGGAGA
Db		
	7118	AAAGGTTACTACTCTCACCGTGACCCACACCCCTTGCACGGCTGCGTGGGAGA
Qy		
	1208	CAGCTAGGCACACTCCAGTCAACTCTCTGGCTAGGCAACATCATGTATGCGCCCACTT
Db		
	7178	CAGCTAGACACACTCCAATCAACTCTTGGCTAGGCAATATCATATGATGCGCCACCC
Qy		
	1268	TGTGGCAAGGATGATCTGATGATCACTTCTTCTCCATCTCTTAGCCAGGAGCAAC
Db		
	7238	TATGGCAAGGATGATCTGATGATCACTTTTCTCATCTCTTAGCTCAAGAGCAAC
Qy		
	1328	TTGAAAAGCCCTGGATTGTTCAGATCTACGGGGCTGTACTCCATTGAGCCACTTGAAC
Db		
	7298	TTGAAAAGCCCTGGATTGTTCAGATCTACGGGGCTTGTCTACTCTCATTTGAGCCACTTGAAC
Qy		
	1388	TACCTCAGATCATTTGAACGACTCCATGGTCTTAGCGCATTTTCACTCCATAGTTACTCTC
Db		
	7358	TACCTCAGATCATTTGAACGACTCCATGGTCTTAGCGCATTTTCACTCCACAGTTACTCTC
Qy		
	1448	CAGGTGAGATCAATAGGTTGGCTTCATGCCCTCAGGAAGCTTGGGTTACCAACCTTGGAG
Db		
	7418	CAGGTGAGATCAATAGGTTGGCTTCATGCCCTCAGGAAGCTTGGGTTACCAACCTTGGAA
Qy		
	1508	TCTGAGACATCGGGCCAGAAGTGTTCGGCGCTAAGTTACTGTCCAGGGGGGAGGGCCG
Db		
	7478	CCTGAGACATCGGGCCAGAAGTGTTCGGCGCTAAGTTACTGTCCAGGGGGGAGGGCCG
Qy		
	1568	CCATTTGTGGCAAGTACCTCTTCAACTGGCAGTAAGGACCAAGCTTAAACTCACTCCAA
Db		
	7538	CCACTGTGTGGCAGATACCTCTTTAACTGGGCAGTAAGGACCAAGCTTAAACTCACTCCAA
Qy		
	1628	TTCCGGCTCGGTCCCGGCTGGACTTGTCCGGCTGGTTCGTGTGCTGAGTACAGCGGGGAG
Db		
	7598	TCCCGGCCCGGTCCCGAGCTGGACTTGTCTGGCTGGTTCGTGTGCTGGTTACAGCGGGGAG
Qy		
	1688	ACATATATCACAGCCTGTCTCGTGGCCGACCCCGCTGGTTCATGTTGTGCTACTCTCTAC
Db		
	7658	ACATATATCACAGCCTGTCTCGTGGCCGACCCCGCTGGTTCATGTTGTGCTACTCTCTAC
Qy		
	1748	TCTCCGTAGGGTAGGCATCTATCTACTCTCCCAACCG
Db		
	7718	TTTCTGTAGGGTAGGCATTTACCTGTCTCCCAACCG

RESULT 9
ADD93730

ADD93730
ID ADD93730 standard: DNA: 7979 BP.

ID ADD93730
XX
AC ADD93730;

XX 29-TAN-2004 (first entry)

XX 25

XX
KW
HCY- vaccine: virucide: ES.

XX	Hepatitis C virus.
OS	
XX	WO2003085084-A2.
XX	
XX	16-OCT-2003.
PD	
XX	
XX	03-APR-2003; 2003WO-US010177.
PF	
XX	
XX	03-APR-2002; 2002US-0369685P.
ER	
XX	(SMIK) SMITHKLINE BEECHAM CORP.
XX	
XX	Gates A, Gu B, Sarisky RT;
PI	
XX	
XX	WPI; 2003-804301/75.
DR	
XX	New hepatitis C virus (HCV) sub-genomic replicon, useful for facilitating
PT	screening or testing of anti-HCV drugs, comprises a nucleic acid
PT	construct encoding chimeric HCV non-structural proteins, and an NS5B
PT	polymerase gene.
PT	

```
Query Match      88.5%; Score 1596.2; DB 10; Length 7979;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 113; Indels 0; Gaps 0
```

Qy	8	GCATGTCAATGTCCTATATCATGTGACAGCGCCCTGTATCACACCGTGTGGCTTGGCGGAGAA	67
Db	5978	GCTGTCTCAATGTCCTATACGTGTGACAGCGCCCTGTATCAGCCATCGCTTGGCGGAGAA	603
Qy	68	GCAAGCTGCCATCAACCGCTGTAGCAACTCCTTTGCTGCGTCAACATAACCTCGTCTATT	127
Db	6038	GTAAGCTGCCATCAACCGCTTTGAGCAACTCTTTGCTGCGTCAACATAACCTCGTCTATT	609
Qy	128	CCACAACATCCCGCAGTCCAAGCCTTCGGCAGAGAAAGGTCACTTTTGACAGACTGCAAG	187
Db	6098	CCACAACATCCCGCAGCGAAGCCTTCGGCAGAGAAAGGTCACTTTTGACAGACTGCAAG	615
Qy	188	TCCTGCACGATCATTTACCGGACGTGTCTCAAGGAGATGAAGGCGAAGCGGTCCACAGTGA	247
Db	6158	TCCTGCAGTGAATCATTTACCGGACGTACTCAAGGAGATGAAGGCGAAGCGGTCCACAGTTA	621
Qy	248	AGGCTTAACCTGCTATCTCTGTAGAAGAAGCATCAAGCTTGACGCCCCCGCATTTGGCCAAAT	307
Db	6218	AGGCTTAAGCTTCTATCTATATAGGAGGCGCTGCAAGCTGAAGCCCCCAACATTCGGCCAAAT	627
Qy	308	CCAAATTTTGGCTATGGGCGAAGAGCGTCCGAGCGCTTCCAGCAGGCGCGCTTTAACACA	367
Db	6278	CCAAATTTGGCTATGGGCGAAGAGCGTCCGAGACCTTATCAGCAGGCGCGCTTTAACACA	633

CC replication. The sub-genomic HCV replicon systems of the invention may
CC provide the foundation for generating HCV replicons of all 6 major
CC genotypes and subtypes to facilitate screening, testing and evaluating
CC anti-infective agents for HCV disease(s).

SQL Sequence 7979 BP; 1673 A; 2351 C; 2216 G; 1739 T; 0 U; 0 Other;

Query Match		88.5%;	Score 1596.2;	DB 10;	Length 7979;
Best Local Similarity		93.6%;	Pred. No. 0;		
Matches 1664;		Conservative 0;	Mismatches 113;	Indels 0;	Gaps 0;
QY	8	GCATGTCAATGTCCTATACATGGACAGCGCCCTGATCACACCGTGGCGTGCAGGAGAA	67		
DB	5978	GCTGCTCAATGTCCTATACGTGGACAGCGCCCTGATCAGCCATGCGCTGCGAGAGAA	6037		
QY	68	GCAAGCTGCCCATCAAGCGCTGAGCAACTCCTTGCTGCGTCCACATAAACCCTGCTATT	127		
DB	6038	GTAAGCTGCCCATCAACCCGTTGAGCACTTTGCTGCGTCCACCAACAATGGTCTACG	6097		
QY	128	CCACAACATCCCGCAGTGCAGAGCTGCGGCAGAGAGAGGTACCTTTGACAGACTGCAAG	187		
DB	6098	CCACAACATCCCGCAGCGCAAGCCTCCGCGAGAGAGAGGTACCTTTGACAGATTGCAAG	6157		
QY	188	TCCTGGAGGATCATTACCGGACGCTGCTCAAGAGATGAAGGCGAAGGCGTCCACAGTGA	247		
DB	6158	TCCTGGAGGATCATTACCGGAGCGTACTCAAGGAGATGAAGGCGAAGGCGTCCACAGTTA	6217		
QY	248	AGGCTAAACTGCTATCTGTAGAAGAAGCATGCAAGCTGAGCCCGCCGCAATTCGGCCAAAT	307		
DB	6218	AGGCTAAAGCTTCTATCTATAGAGAGGCGCTGCAAGCTGAGCCCGCCACATTCGGCCAAAT	6277		
QY	308	CCAAATTTGGCTATGGGCGAAAGACGTCGGAGCCTATCAGCAGGCGCGTTAAACCA	367		
DB	6278	CCAAATTTGGCTATGGGCGAAAGACGTCGGAGCCTATCAGCAGGCGCGTTAAACCA	6337		
QY	368	TCCGCTCGGTGGAAGGACTTGTGAGGACACTGACACACCAATTCAGACCAACATCA	427		
DB	6338	TCCGCTCGGTGGAAGGACTTGTGGAAGACACTGAAACACCAATTCAGACCAACATCA	6397		
QY	428	TGGCAAAAATGAGGTTTCTGCTCCACAGAGAAAGGAGCGCGCAACCCAGCTGCC	487		
DB	6398	TGGCAAAAATGAGGTTTCTGCTCCACAGAGAAAGGAGCGCGCAACCCAGCTGCC	6457		
QY	488	TCACTGTAATCCAGACCTGGAGTTTCTGTATCGGAGAGATGGCCCTTACGACGTGG	547		
DB	6458	TATATGTAATCCAGACCTGGAGTTTCTGTATCGGAGAGATGGCCCTTACGACGTGG	6517		
QY	548	TTTCCACTCTTCTCAGGCGGTGATGGCTCTCATACGGATTCCAATACTCTCTAAGC	607		
DB	6518	TCTCCACCCTTCTCAGGCGGTGATGGCTCTCATACGGATTCCAATACTCTCTAAGC	6577		
QY	608	AGCGGTGCGATTCTGCTGTAATACCTGGAAGCAAGAAATGCCCTATGCGCTTCTCAT	667		
DB	6578	AGCGGTGCGATTCTGCTGTAATACCTGGAATCAAAAGAAATGCGCTATGCGCTTCTCAT	6637		
QY	668	ATGACACCGCTGTTTTCACCTCAACGGTCACTGAGATGACATCCGTGTTGAGGAGTCAA	727		
DB	6638	ATGACACCGCTGTTTTCACCTCAACGGTCACTGAGAGTGAATTCGTGTTGAGGAGTCAA	6697		
QY	728	TTTACCAATGTTTGACTTTGGCCCCCGAAGCTAGACAGAGCCCATTAAGTCTGCTCACAGAGC	787		
DB	6698	TTTACCAATGTTTGACTTTGGCCCCCGAAGCTAGACAGAGCCCATTAAGTCTGCTCACAGAGC	6757		
QY	788	GCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGCGGCTATCCCGGT	847		
DB	6758	GGCTTTACATCGGGGTCCCGTGACTAACTCAAAAGGCGAGAACTGCGGCTTATCGCGGT	6817		
QY	848	GCCGCGCAGCGGCTGCTGAGACTAGCTCGGTAAATACCTCACATGCTACTTGAAGG	907		
DB	6818	GCCGCGCAAGTGGCGTGTGACACTAGCTCGGTAAATACCTCACATGTTTACTTGAAGG	6877		
QY	908	CGCTGACAGCTGTCGAGCTCCAAAGCTCCAGGACTGACCATGCTGTTGAATGAGAGC	967		

RESULT 11

ADD93731

ID ADD93731 standard; DNA; 7979 BP.

XX ADD93731;

XX 29-JAN-2004 (first entry)

XX Hepatitis C virus strain J4 replicon.

XX HCV; vaccine; virucide; ss.

XX Hepatitis C virus; genotype 1b.

DB	6878	CCACTGACGCTGTGAGCTGCAAAAGCTCCAGGACTGCACGATGCTCTGTAACGAGAGCG	6937		
QY	968	ACCTTGTGCTTATCTGTGAAAGCGGGGAAACCCAGAGGACGCGGCAAGCTTACAGTCT	1027		
DB	6938	ACCTTGTGCTTATCTGTGAAAGCGGGGAAACCCAGAGGATGCGGCGGCTTACGAGCT	6997		
QY	1028	TCACGAGGCTATGACTAGGTACTCTGCCCGCTTGGGACCGCCCGCCCAACCGGAATACG	1087		
DB	6998	TCACGAGGCTATGACTAGGTACTCTGCCCGCTTGGGACCGCCCGCCCAACCGGAATACG	7057		
QY	1088	ACTTGGAGCTGATAACATCGTGTCTTCCAATGTGTGCGTTCGACACAGATGCATCTGGCA	1147		
DB	7058	ACTTGGAGCTGATAACATCATGTTCTTCCAATGTGTGAGTGGCGCACGATGCATCTGGCA	7117		
QY	1148	AAAGGTGTACTTACTTACCTGACCCGTCACCGTACCGTCCCTTGGGACCGCCCGCCCAACCGGAATACG	1207		
DB	7118	AAAGGTGTACTTACTTACCTGACCCGTCACCGTACCGTCCCTTGGGACCGCCCGCCCAACCGGAATACG	7177		
QY	1208	CAGCTAGGACAGACTCCAGTCAACTCTGCTAGGCAAGCATCATGATGCGCCGCTT	1267		
DB	7178	CAGCTAGGACAGACTCCAGTCAACTCTGCTAGGCAAGCATCATGATGCGCCGCTT	7237		
QY	1268	TGTGGCAAGGATGATTCTGATGACTCACTTCTTCTTCCATCTTCTAGCCAGGAGCAAC	1327		
DB	7238	TATGGCAAGGATGATTCTGATGACTCACTTCTTCTTCCATCTTCTAGCTCAAGAGCAAC	7297		
QY	1328	TTGAAAAGCCCTGGATTGTGAGATCTACGGGCTTGTCTACTTCCATTTAGCCACTTGACC	1387		
DB	7298	TTGAAAAGCCCTGGATTGTGAGATCTACGGGCTTGTCTACTTCCATTTAGCCACTTGACC	7357		
QY	1388	TACTCAGATCATTTGAAACGACTCCATGCTTATAGGCGATTTTACATCCATAGTACTCTC	1447		
DB	7358	TACTCAGATCATTTGAAACGACTCCATGCTTATAGGCGATTTTACATCCATAGTACTCTC	7417		
QY	1448	CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAGCTTGGGTTACCCACCTTGGAG	1507		
DB	7418	CAGGTGAGATCAATAGGTGGCTTCATGCTCAGGAGCTTGGGTTACCCACCTTGGAG	7477		
QY	1508	TCTGGAGACATCGGGGCGAGAGTGTCCGCGCTTAACTTACTGTCCAGGGGGGAGGGCGG	1567		
DB	7478	CCTGGAGACATCGGGGCGAGAGTGTCCGCGCTTAACTTACTGTCCAGGGGGGAGGGCGG	7537		
QY	1568	CCATTTGTGCGAGTACCTCTTCAACTGGGCGAGTAAAGCAAGCTTAAACTCACTCCAA	1627		
DB	7538	CCATTTGTGCGAGTACCTCTTCAACTGGGCGAGTAAAGCAAGCTTAAACTCACTCCAA	7597		
QY	1628	TTCCGCTCGTCCCGGCTGGAGTGTGCTCGGCTGGTTCGTTGCTGCTACAGGGGGAG	1687		
DB	7598	TTCCGCTCGTCCCGGCTGGAGTGTGCTCGGCTGGTTCGTTGCTGCTACAGGGGGAG	7657		
QY	1688	ACATATATCACAGCTGCTCTGTCGCGCGACCCCGCTGCTTCACTTGTGCTACTCTTAC	1747		
DB	7658	ACATATATCACAGCTGCTCTGTCGCGCGACCCCGCTGCTTCACTTGTGCTACTCTTAC	7717		
QY	1748	TCTCCGTAGGGGTAGGATCTATCTACTCCCAACCG 1784			
DB	7718	TTTCTGTAGGGGTAGGATCTATCTACTCCCAACCG 7754			


```
OY 848 GCCGCGCAGCGCGCTGCTGAGCACTAGCTGGGTAAATACCTCACAATGCTACTTGAAGG 907
Db      |||
OY 8434 GCGCGCAAGTGGCTGCTGAGCACTAGCTGGGTAAATACCTCACAATGCTACTTGAAGG 8493
Db      |||
OY 908 CCGCTGCACGCTGTGAGAGTCCAAAGCTCCAGAGCTGCACGATGCTCGTGAATGGAGAGG 967
Db      |||
OY 8494 CCACTGCACGCTGTGAGAGTCCAAAGCTCCAGAGCTGCACGATGCTCGTGAATGGAGAGG 8553
Db      |||
OY 968 ACCTTGTCTTATCTGTGAAGCGGGAACCAAGAGAGCAAGCGGAGCTTACAGTCT 1027
Db      |||
OY 8554 ACCTTGTCTTATCTGTGAAGCGGGAACCAAGAGAGCAAGCGGAGCTTACAGTCT 8613
Db      |||
OY 1028 TCACGAGGCTATGACTAGTACTCTGCCCCCTCGGGGACCGCCCCCAACCGGATACG 1087
Db      |||
OY 8614 TCACGAGGCTATGACTAGTACTCTGCCCCCTCGGGGATCCGCCCAACAGATACG 8673
Db      |||
OY 1088 ACTTGAGCTGATAACATCGTGTCTCCCAATGTGTGCTGCGACACACATGATCTGGCA 1147
Db      |||
OY 8674 ACTTGAGCTGATAACATCGTGTCTCCCAATGTGTGCTGCGACACATGATCTGGCA 8733
Db      |||
OY 1148 AAGGGTGTACTCTACCTACCGGTGACCCACCGTCCCCCTTGGCGGGCTGCTGGGAGA 1207
Db      |||
OY 8734 AAGGGTGTACTCTACCTACCGGTGACCCACCGTCCCCCTTGGCGGGCTGCTGGGAGA 8793
Db      |||
OY 1208 CAGCTAGGCACTCCAGTCACTCTGCTAGGCAACATCATCATGATGCGGCCACTT 1267
Db      |||
OY 8794 CAGCTAGGCACTCCAGTCACTCTGCTAGGCAACATCATCATGATGCGGCCACTT 8853
Db      |||
OY 1268 TGTGGCAAGGATGATCTGATGACTCACTTCTCTCCATCTTACCCAGGAGCAAC 1327
Db      |||
OY 8854 TATGGCAAGGATGATCTGATGACTCACTTCTCTCCATCTTACCCAGGAGCAAC 8913
Db      |||
OY 1328 TTGAAAGCCCTTGGATTGTGATGACTCTACGGGCTTGTACTCCATTTAGCCACTTGACC 1387
Db      |||
OY 8914 TTGAAAGCCCTTGGATTGTGATGACTCTACGGGCTTGTACTCCATTTAGCCACTTGACC 8973
Db      |||
OY 1388 TACTCTAGATCATGAAAGCTCCATGCTTTAGCGCAATTTTCACTCCATAGTTACTCTC 1447
Db      |||
OY 8974 TACTCTAGATCATGAAAGCTCCATGCTTTAGCGCAATTTTCACTCCATAGTTACTCTC 9033
Db      |||
OY 1448 CAGGTGAGATCAATAGGTGGCTTATGCTCAGGAAGCTTGGGTACACCCCTTGGAG 1507
Db      |||
OY 9034 CAGGTGAGATCAATAGGTGGCTTATGCTCAGGAAGCTTGGGTACACCCCTTGGAG 9093
Db      |||
OY 1508 TCTGAGACATCGGCGCAGAGTGTCCGGCTAAGTGTACTGTCAGGGGGGGGGGGCG 1567
Db      |||
OY 9094 CTTGAGACATCGGCGCAGAGTGTCCGGCTAAGTGTACTGTCAGGGGGGGGGGGCG 9153
Db      |||
OY 1568 CCATTTGTGGCAAGTACCTCTTCAACTGGGCAAGTAAAGCAACCAAGCTTAAACTCCAA 1627
Db      |||
OY 9154 CCATTTGTGGCAAGTACCTCTTCAACTGGGCAAGTAAAGCAACCAAGCTTAAACTCCAA 9213
Db      |||
OY 1628 TTCCGGCTGCGTCCGGCTGAGCTTGTGCGGTGCTTGTGCTGCTACAGCGGGGAG 1687
Db      |||
OY 9214 TCCCGGCGCGTCCAGCTGAGCTTGTGCTGCGGTGCTTGTGCTGCTTACAGCGGGGAG 9273
Db      |||
OY 1688 ACATATATACAGCTGTCTGCTGCGGACCCCGCTGCTTCAATGTTGCTTACTCTCTAC 1747
Db      |||
OY 9274 ACATATATACAGCTGTCTGCTGCGGACCCCGCTGCTTCAATGTTGCTTACTCTCTAC 9333
Db      |||
OY 1748 TCTCGTGGGGTAGGCTATCTACTCTCCCAACCG 1784
Db      |||
OY 9334 TTTCTGTAGGGTAGGCTATTTACTCTCTCCCAACCG 9370
Db      |||
```

RESULT 13

AAF23492

ID AAF23492 standard; DNA; 9595 BP.

XX AC

AAF23492;

XX AC

DT 21-MAR-2001 (first entry)

XX XX

DE Infectious Hepatitis C virus 1b genotype.

XX GBV-B; hepatitis C virus; HCV; vaccine; ds.

XX Hepatitis C virus.

OS OS

XX WO200075337-A1.

PN 14-DEC-2000.

XX 02-JUN-2000; 2000WO-US015293.

XX 04-JUN-1999; 99US-0137694P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Bukh J, Yanagi M, Emerson SU, Purcell RH;

XX MPI; 2001-091214/10.

XX New infectious nucleic acids of the GB virus-B clone, useful for

PT indirectly studying the molecular properties of hepatitis C virus (HCV)

PT and in developing vaccines and therapeutics for HCV.

XX Disclosure; Fig 7; 96pp; English.

XX The present invention relates to GB virus-B. The nucleic acid molecules

CC of the invention are useful for indirectly studying the molecular

CC properties of hepatitis C virus (HCV). The infectious nucleic acid

CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used

CC in the development of vaccines and therapeutics for HCV

XX SQ

Sequence 9595 BP; 1934 A; 2843 C; 2697 G; 2121 T; 0 U; 0 Other;

Query Match 88.5%; Score 1596.2; DB 4; Length 9595;

Best Local Similarity 93.6%; Pred. No. 0;

Matches 1664; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

OY 8 GCATGTCAATGCTCTATACATGGACAGCGCCCTGTGATCAGCCATGCGTCCGGAGGAAA 67

Db 7594 GCTGCTCAATGCTCTATACATGGACAGCGCCCTGTGATCAGCCATGCGTCCGGAGGAAA 7653

OY 68 GCAAGCTGCCCATCAACGCGCTGAGCAACTCTTGTGCTGCGTCACCAATACCTGCTCTATT 127

Db 7654 GTAGCTGCCCATCAACGCGCTGAGCAACTCTTGTGCTGCGTCACCAATACCTGCTCTAG 7713

OY 128 CCACAAATCCCGCAGTGCAGGCTCGGCGAGAGAGGTCCCTTTGACAGACTGCAAG 187

Db 7714 CCACAAATCCCGCAGGCGCAAGCCTCCGCGAGAGAGGTCCCTTTGACAGACTGCAAG 7773

OY 188 TCCTGGAGGATCATTTACGGGACGCTGCTCAAGGAGATGAAGGCGGAGCGTCCACAGTGA 247

Db 7774 TCCTGGAGGATCATTTACGGGACGCTGCTCAAGGAGATGAAGGCGGAGCGTCCACAGTGA 7833

OY 248 AGGTAAACTGCTATCTGTAGAGAGCATGCAAGCTGAGCGCCCGCCCGCATTTGCGCCAAAT 307

Db 7834 AGGTAAACTGCTATCTGTAGAGAGCATGCAAGCTGAGCGCCCGCCCGCATTTGCGCCAAAT 7893

OY 308 CCAATTTGGCTATGGGCGCAAGGACCTCCGGAGCGCTTCCAGCGGCGCGCTTAAACCA 367

Db 7894 CCAATTTGGCTATGGGCGCAAGGACCTCCGGAGCGCTTCCAGCGGCGCGCTTAAACCA 7953

OY 368 TCCTGCTCCGTTGGAAGGACTTGTGAGGAGACTGACACCAATTTGACAGCACTCA 427

Db 7954 TCCTGCTCCGTTGGAAGGACTTGTGAGGAGACTGACACCAATTTGACAGCACTCA 8013

OY 428 TGGCAAAAATGAGTTTCTGCTGCCAACCAGAGAGGAGCGCGCAACACCACTGCGC 487

Db 8014 TGGCAAAAATGAGTTTCTGCTGCCAACCAGAGAGGAGCGCGCAACACCACTGCGC 8073

OY 488 TCATCGTATTTCCCAAGCTGCGAGTTCGTGTATCGGAGAGAGTGGCCCTTTACGACGTTGG 547

Db 8074 TATCGTATTTCCCAAGCTGCGAGTTCGTGTATCGGAGAGAGTGGCCCTTTACGACGTTGG 8133

Best Local Similarity 93.6%; Pred. No. 0;			
Matches 1664; Conservative 0; Mismatches 113; Indels 0; Gaps 0;			
QY	8	GCATGTCATATGCTTATACATGACAGGCGCCCTGATCACACCGTGGCTGCGAGGAAA	67
DB	7594	GCTGCTCAATGTCCTATATGATGACAGGCGCCCTGATCACCGCCATGCGCTGCGAGGAAA	7653
QY	68	GCAAGCTGCCATCAACGCGCTGAGCAACTCTTGTGGTGCACATTAACCTGGTCTATT	127
DB	7654	GTAAGCTGCCATCAACCGCTGTAGCAACTCTTGTGGTGCACATTAACCTGGTCTAG	7713
QY	128	CCACAACATCCCGCATGCAAGCTGCGGCAGAGAAGGTCACTTTGACAGACTGCAAG	187
DB	7714	CCACAACATCCCGCAGCGCAAGCTTCCGCGAGAAGGTCACTTTGACAGATTGCAAG	7773
QY	188	TCCTGGACGATCAATACCGGGACGTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTGA	247
DB	7774	TCCTGGATGATCAATACCGGGACGTACTCAAGGAGATGAAGCGGAGCGTCCACAGTTA	7833
QY	248	AGGCTAACTGCTATCTGTAGAGAAGCATGCAAGCTGACGCCCGGCAATTGCGGCAAT	307
DB	7834	AGGCTAAGCTTCTATCTATAGAGGAGCGCTGCAAGCTGACGCCCGGCAATTGCGGCAAT	7893
QY	308	CCAAATTTGGCTATGGGCAAGGAGCTCCGAGGCTATCCAGAGGCGGCTTAACCA	367
DB	7894	CCAAATTTGGCTATGGGCAAGGAGCTCCGGAACCTATCCAGAGGCGGCTTAACCA	7953
QY	368	TCGCTCCGTGTGGAGGACTTCTGGAGGACACTGACACACCAATTCAGACCCATCA	427
DB	7954	TCGCTCCGTGTGGAGGACTTCTGGAGGACACTGAAACCAATTCAGACCCATCA	8013
QY	428	TGGCAAAAATGAGTGTCTCGCTCCAAACAGAGAAAGCGCGCAAAACGAGCTCGCC	487
DB	8014	TGGCAAAAATGAGTGTCTCGCTCCAAACAGAGAAAGCGCGCAAAACGAGCTCGCC	8073
QY	488	TCATCGTATTCACAGCTGGAGTTGCTGTATGCGAAGATGCCCTTTACGACGTGG	547
DB	8074	TTATCGTATTCACAGCTGGAGTTGCTGTATGCGAAGATGCCCTTTACGACGTGG	8133
QY	548	TTTCCACTTCTCCTCAGGCGGTGATGGGCTCCTCATACGGATTCCAATCTCTCTAAGC	607
DB	8134	TTTCCACTTCTCCTCAGGCGGTGATGGGCTCCTCATACGGATTCCAATCTCTCTAAGC	8193
QY	608	AGCGGTGAGTTCCTGGTGAATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCAT	667
DB	8194	AGCGGTGAGTTCCTGGTGAATACCTGGAATCAAGAAATGCCCTATGGGCTTCTCAT	8253
QY	668	ATGACACCGCTGTGTTGACTCAACGCTCAGTCAATGACATCGCTGTGAGAGTCAA	727
DB	8254	ATGACACCGCTGTGTTGACTCAACGCTCAGTCAATGACATCGCTGTGAGAGTCAA	8313
QY	728	TTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCCATAAGGTGCTTCACAGAGC	787
DB	8314	TTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCCATAAGGTGCTTCACAGAGC	8373
QY	788	GGCTCTATGTGGGGGTCCCATGACTAATCAAGAGGAGAACTGGCGGCTATGCGCGGT	847
DB	8374	GGCTTTATACCGGGGTCCCATGACTAATCAAGAGGAGAACTGGCGGCTATGCGCGGT	8433
QY	848	GCGCGCAGCGGCTGTGACGACTAGTGGGTAAATACCTCACATGCTACTTTGAAGG	907
DB	8434	GCGCGCAGCGGCTGTGACGACTAGTGGGTAAATACCTCACATGCTACTTTGAAGG	8493
QY	908	CCGCTGCAGCCTGTGAGCTGCAAGCTCCAGGACTGCAAGATGCTCGTGAATGGAGACG	967
DB	8494	CCATGCGAGCCTGTGAGCTGCAAGCTCCAGGACTGCAAGATGCTCGTGAACGAGACG	8553
QY	968	ACCTTGTGTTATCTGTGAAGCGCGGAAACCAAGAGGACCGGCAAGCTACAGTCT	1027
DB	8554	ACCTTGTGTTATCTGTGAAGCGCGGAAACCAAGAGGATGCGGCGGCCCTACAGCT	8613
QY	1028	TCACGAGGCTATGACTAGTACTCTGCCCGCTTGGGAGCGCCGCCCAACCGGATACG	1087

DB	8614	TCACGAGGCTATGACTAGGTATTTCCGCCCCCCCCGGGATCCGCCCAACAGATACG	8673
QY	1088	ACTTGGAGCTGATAACATCGTGTCTCTCCAATGTTCGGTGCACACAGATGATCGCA	1147
DB	8674	ACCTGGAGCTGATAACATCATGTTCTCTCCAATGTTCAGTTCGGCAAGATGATCGCA	8733
QY	1148	AAAGGGTGTACTCTCTACCCGTGACCCACCGTCCCTTGGCGGGCTGCTGGGAGA	1207
DB	8734	AAAGGGTATACTACTCTACCCGTGACCCACCGTCCCTTGGCAAGGCTGCTGGGAGA	8793
QY	1208	CAGTAGGACACATCCAGTCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1267
DB	8794	CAGTAGGACACATCCCAATCAACTCTTGGCTAGGCAATATCATATGATGGGCCACCC	8853
QY	1268	TGTGGCAAGAGATGATCTGATGACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT	1327
DB	8854	TATGGCAGAGATGATCTGATGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	8913
QY	1328	TTGAAAAAGCCCTGGATGTCAGATCTAAGGGCTTGTATTCTCATTTAGGACCACTGACC	1387
DB	8914	TTGAAAAAGCCCTGGATGTCAGATCTAAGGGCTTGTATTCTCATTTAGGACCACTGACC	8973
QY	1388	TACCTCAGATCATTTGAACGACTCCATGCTCTTAGGGCATTTTCACTCCATAGTTACTCTC	1447
DB	8974	TACCTCAGATCATTTGAACGACTCCATGCTCTTAGGGCATTTTCACTCCATAGTTACTCTC	9033
QY	1448	CAGGTGAGATCAATAGGGTGGCTTCATGCCCTCAGGAAGCTTGGGGTACCACTTGGCAA	1507
DB	9034	CAGGTGAGATCAATAGGGTGGCTTCATGCCCTCAGGAAGCTTGGGGTACCACTTGGCAA	9093
QY	1508	TCTGGAGACATCGGGCCAGAGTGTCCGCGCTAAGTTACTGTCCAGGGGGAGGGCGG	1567
DB	9094	CCTGGAGACATCGGGCCAGAGTGTCCGCGCTAAGTTACTGTCCAGGGGGAGGGCGG	9153
QY	1568	CCATTGTGGCAAGTACTCTTTCAACTGGGAGTAAGGACCAAGCTTAAACTCACTCCAA	1627
DB	9154	CCACTTGTGGCAAGTACTCTTTAACTGGGAGTAAGGACCAAGCTTAAACTCACTCCAA	9213
QY	1628	TTCCGGCTGGTCCGGCTGAGTGTCCGGCTGGTGTGGTGTGGTGTGGTGTGGTGTGG	1687
DB	9214	TTCCGGCTGGTCCGGCTGAGTGTCCGGCTGGTGTGGTGTGGTGTGGTGTGGTGTGG	9273
QY	1688	ACATATATCAGAGCTGTCTGCTGCCGACCCCGCTGGTTCATGTTGGCTACTCTCTAC	1747
DB	9274	ACATATATCAGAGCTGTCTGCTGCCGACCCCGCTGGTTCATGTTGGCTACTCTCTAC	9333
QY	1748	TCTCCGTAGGGGTAGGCATCTATCTACTTCCCCCAACCG 1784	
DB	9334	TTTCTGTAGGGGTAGGCATTTACTGTCTCCCAACCG 9370	

RESULT 15

AD036222

ID AD036222 standard; DNA; 9595 BP.

XX AD036222;

AC AD036222;

DT 26-AUG-2004 (first entry)

XX Hepatitis C virus (HCV) J416 wild-type cDNA.

XX hepatotropic; virucide; vaccine; gene therapy; vaccine;

KW Hepatitis C virus; HCV; core protein; HCV infection; vaccination; gene;

XX Hepatitis C virus.

XX Key Location/Qualifiers

FH 342..9374

FT /*tag= a

FT /product= "HCV polypotein"

XX MO2004046175-A1.

XX 03-JUN-2004.
PD 13-NOV-2003; 2003WO-EP012793.
XX 15-NOV-2002; 2002GB-00026722.
XX (GLAX) GLAXO GROUP LTD.
XX Brett S, Hamblin PA, Ogilvie L;
PI WPI; 2004-420613/39.
DR P-PSDB; ADO36227.
DR
XX New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT the polypeptide sequences of the HCV core and at least one other HCV
PT protein, for use in medicine, particularly for manufacturing a medicament
PT for treating HCV.
XX
PS Disclosure; Fig 1; 78pp; English.
XX
CC The invention describes a polynucleotide vaccine comprising a
CC polynucleotide sequence (S1) encoding the Hepatitis C virus (HCV) Core
CC protein and at least 1 other HCV protein, and causes expression of the
CC proteins in cells (in which (S1) has been mutated or positioned relative
CC to the polynucleotide sequence encoding the other HCV protein, so that
CC the negative effect of the Core protein on expression of the other HCV
CC protein is reduced). Also described are: a method of preventing or
CC treating an HCV infection in a mammal, comprising administering the
CC vaccine cited above to a mammal; and a method of vaccination of an
CC individual, comprising taking a polynucleotide vaccine as cited above,
CC coating the polynucleotide onto gold beads and delivering the gold beads
CC into the skin. HCV nucleic acids, polypeptides, host cells, vectors and
CC antibodies used in the methods, are also disclosed. The polynucleotide
CC vaccine is useful in the manufacture of a medicament for the treatment of
CC HCV. This sequence encodes the wild type HCV polypeptide.
XX
SQ Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;

Query Match 88.5%; Score 1596.2; DB 12; Length 9595;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 8 GCATGTCATGTCCTATACATGGACAGCGCCCTGATCACACCGTGGCGTGGGAGGAAA 67
DB 7594 GCTGCTCAATGTCTTATACGTGGACAGCGCCCTGATCACGCCATGGCTGGGAGGAAA 7653

QY 68 GCAAGCTGCCATCAACGCGGTGAGCAACTCCTTGCTGCGTCAACATAACCTGGTCTATT 127
DB 7654 GTAAGCTGCCATCAACCGGTGAGCAACTCTTGCTGCGTCAACATAACCTGGTCTACG 7713

QY 128 CCACAACATCCCGCAGTGCAGCCCTGGCGGAGAGAGAGGTCAACCTTTGACAGACTGCAAG 187
DB 7714 CCACAACATCCCGCAGCGCAAGCCTCCGCGAGAGAGAGGTCAACCTTTGACAGATTGCAAG 7773

QY 188 TCCTGGAGATCAITATACCGGACAGTGTCAAGAGATGAAGGCGAGAGCGTCCACAGTGA 247
DB 7774 TCCTGGATGATCATTTACCGGGACGTACTCAAGGAGATGAAGGCGAGAGCGTCCACAGTTA 7833

QY 248 AGGCTAAACTGCTATCTGTAGAAGBAGCATGCAAGCTGACGCCGCCCGCATTCGGGCCAAAT 307
DB 7834 AGGCTAAAGCTTCTATCTATPAGAGGAGCGCTGCAAGCTGACGCCGCCCGCATTCGGGCCAAAT 7893

QY 308 CCAAAATTTGGCTATCGGGCAAGGACGCTCCGAGCGCTATCCAGCAGAGGCGCGTTAAACCACA 367
DB 7894 CCAAAATTTGGCTATCGGGCAAGGACGCTCCGAGACCTATCCAGCAGAGGCGCGTTAAACCACA 7953

QY 368 TCCGCTCCGCTGGAGGACTTGTGGAGGACACTTGACACACCAATTCAGACACCAATCA 427
DB 7954 TCCGCTCCGCTGGAGGACTTGTGGAGGACACTTGAAGACACTTGAACCAATTCAGACACCAATCA 8013

QY 428 TGGCAAAAATAGGTTTTCGCGTCCAAACAGAGAAAGGAGGCGCGCAACAGCTGCC 487

DB 8014 TGGCAAAAAGTGGGTTTTTCTCGGTCCAAACACAGAGAGGAGCGGCCAGCTCGCC 8073

QY 488 TCATCGTATTTCCAGACCTGGGAGTTGCTGTATGCGAGAAAGATGGCCCTTTTACGAGCTGG 547
DB 8074 TTATCGTATTTCCAGACCTGGGAGTTGCTGTATGCGAGAAAGATGGCCCTTTTACGAGCTGG 8133

QY 548 TTTCCACTCTTCTCAGGCGGTGATGGGCTCTCTACATCGGATTCGAATACCTCTCTAAAGC 607
DB 8134 TCTCCACCCCTTCTCAGGCGGTGATGGGCTCTCTACATCGGATTCGAATACCTCTCTAAAGC 8193

QY 608 AGCGGTCGAGTTCTCTGGTGAATACCTGCAAGAGCAAGAAATGCCCCATATGSGCTTCTCAT 667
DB 8194 AGCGGTCGAGTTCTCTGGTGAATACCTGGAATCAAGAAATGCCCCATATGSGCTTCTCAT 8253

QY 668 ATGACACCCGCTGTTTTTGACTCAACCGGTCACTGAGAAATGACATTCCTGTGTGAGGAGTCAA 727
DB 8254 ATGACACCCGCTGTTTTGACTCAACCGGTCACTGAGAGTGACATTCGTGTGTGAGGAGTCAA 8313

QY 728 TTTACCAATGTTGTGACTTTGGCCCCCGAAGCTAGACAGGCCATAGGTCGCTCACAGAGC 787
DB 8314 TTTACCAATGTTGTGACTTTGGCCCCCGAGGCCACAGACAGGCCATAGGTCGCTCACAGAGC 8373

QY 788 GGTCTATGTGCGGGGTCCCATCACTAACTCCAAAGGCGAGACTGCGGCTATCGCCGGT 847
DB 8374 GGTCTTACATTCGCGGGTCCCTGACTAACTCAAAAGGCGAGAACTGCGGTATATCGCCGGT 8433

QY 848 GCGCGCGAGCGCGTGTGACGACTAGCTGCGGTAAATACCTCCATCATGCTACTTCTGAAGG 907
DB 8434 GCGCGCAAGTGGCTGTGACGACTAGCTGCGGTAAATACCTCCATCATGCTACTTCTGAAGG 8493

QY 908 CCCTCTCAGCCTGTGAGCTGCCAAGCTCCAGAGCTGCACAGATGCTGTAATGAGAGAGC 967
DB 8494 CCACTGCGCCTGTGAGCTGCCAAGCTCCAGAGCTGCACAGATGCTGTAATGAGAGAGC 8553

QY 968 ACCTTGCTGTATCTGTGAAAGCGCGGAAACCCAGAGAGACGCGGCAAGCTACGAGTCT 1027
DB 8554 ACCTTGCTGTATCTGTGAAAGCGCGGAAACCCAGAGAGATGCGGCGGCGCTACGAGCCT 8613

QY 1028 TCACGGAGGCTATGACTAGGTACTCTGCCCCCCCCCTGGGAGACCGCCCCCAACCGGAATACG 1087
DB 8614 TCACGGAGGCTATGACTAGGTATCTCGCCCCCCCCCGGGGATCCGCCCAACCAAGAAATACG 8673

QY 1088 ACTTGGAGCTGATAACATCGTGTCTTCCAAATGTGTGCGGTGCGCACACGATGCTATCGGCA 1147
DB 8674 ACTTGGAGCTGATAACATCATGTTCCTCCNAATGTGTCACTGGCGACGATGCATCTGGCA 8733

QY 1148 AAAGGCTACTACTCACTCAACCGGTGACCCCAACCGTCCCTTGGCGGGGTGCGTGGGAGA 1207
DB 8734 AAAGGCTACTACTCACTCAACCGGTGACCCCAACCGTCCCTTGGCACGGGTGCGTGGGAGA 8793

QY 1208 CAGCTAGGCGACTTCAGCTCAACTCTGCTAGGCAACATCATCATGTATGCGCCACTT 1267
DB 8794 CAGCTAGACACTTCCAATCAACTCTTGGCTAGGCAATATCATCATGTATGCGCCCAACCC 8853

QY 1268 TGTGGGCAAGGATGTTCTGATGACTCACTTCTTCTCCATCTCTTCTAGCCAGGAGCAAC 1327
DB 8854 TATGGGCAAGGATGATTCGATGACTCACTTTTCTCCATCTCTTCTAGCTCAAGAGCAAC 8913

QY 1328 TTGAAAAAGCCCTGGGATTTGTACAGATCTAGCGGGCTTTGTTACTCTCATTTGAGGCCACTTGACC 1387
DB 8914 TTGAAAAAGCCCTGGATTTGTACAGATCTACGGGGCTTGTACTCTCATTTGAGGCCACTTGACC 8973

QY 1388 TACTCAGATCATTTGAAAGCACTCCATGGTCTTAGCGGATTTTTCAGTCCATAGTACTCTC 1447
DB 8974 TACTCAGATCATTTGAAAGCACTCCATGGTCTTAGCGGATTTTACACTCCACAGTTACTCTC 9033

QY 1448 CAGGTGAGATCAATAGGTTGGCTTCATGCTCAGGAGCTTTGGGGTACCACCTTCGCGAG 1507
DB 9034 CAGGTGAGATCAATAGGTTGGCTTCATGCTCAGGAGCTTTGGGGTACCACCTTCGCGAA 9093

QY 1508 TCTTGGAGACATTCGGGCGCAGAAAGTGTCCGCGCTAAGTTACTGTGCCAGGGGGGAGGGCCG 1567
DB 9094 CCTTGGAGACATTCGGGCGCAGAAAGTGTCCGCGCTAAGTACTGTGCCAGGGGGGAGGGCCG 9153

This Page Blank (uspto)

GenCore version 5.1.16
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 14:01:52 ; Search time 1271.21 Seconds
(without alignments)
9793.114 Million cell updates/sec

Title: US-10-712-479-1

Perfect score: 1803

Sequence: 1 atggtagcatgtcaatgtc.....ggcatcaccatcaccatcac 1803

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1803	100.0	1803	US-10-712-479-1
2	1722	95.5	1740	US-10-712-479-3
3	1607.4	89.2	1884	US-09-838-386-11
4	1597.8	88.6	5955	US-10-492-178-5
5	1592.8	88.3	37090	US-10-492-178-4
6	1591.4	88.3	5965	US-10-492-178-2
7	1585	87.9	6189	US-10-259-275-41

8	1585	87.9	6189	24	US-11-006-313-41	Sequence 41, Appl
9	1585	87.9	7989	17	US-10-434-842-16	Sequence 16, Appl
10	1585	87.9	7989	19	US-10-639-150-1	Sequence 1, Appl
11	1585	87.9	7989	21	US-10-897-648-17	Sequence 17, Appl
12	1585	87.9	7992	13	US-10-005-469-1	Sequence 1, Appl
13	1585	87.9	7992	13	US-10-005-469-2	Sequence 2, Appl
14	1585	87.9	7992	13	US-10-005-469-4	Sequence 4, Appl
15	1585	87.9	7992	13	US-10-005-469-5	Sequence 5, Appl
16	1585	87.9	7992	13	US-10-005-469-6	Sequence 6, Appl
17	1585	87.9	7992	17	US-10-434-842-1	Sequence 1, Appl
18	1585	87.9	7992	17	US-10-434-842-2	Sequence 2, Appl
19	1585	87.9	7992	17	US-10-434-842-4	Sequence 4, Appl
20	1585	87.9	7992	17	US-10-434-842-5	Sequence 5, Appl
21	1585	87.9	7992	17	US-10-434-842-6	Sequence 6, Appl
22	1585	87.9	7992	17	US-10-434-842-15	Sequence 15, Appl
23	1585	87.9	7995	13	US-10-005-469-3	Sequence 3, Appl
24	1585	87.9	7995	17	US-10-434-842-3	Sequence 3, Appl
25	1585	87.9	9605	16	US-10-467-000-2	Sequence 2, Appl
26	1585	87.9	10690	14	US-10-125-940-1	Sequence 1, Appl
27	1585	87.9	10690	16	US-10-125-920-1	Sequence 1, Appl
28	1585	87.9	10690	18	US-10-467-000-3	Sequence 3, Appl
29	1585	87.9	12305	20	US-10-422-323A-2	Sequence 2, Appl
30	1585	87.9	12315	20	US-10-422-323A-1	Sequence 1, Appl
31	1584.8	87.9	1772	19	US-10-384-339C-132	Sequence 132, App
32	1583.4	87.8	7992	17	US-10-434-842-17	Sequence 17, Appl
33	1583.4	87.8	8638	13	US-10-029-907-6	Sequence 6, Appl
34	1583.4	87.8	8638	13	US-10-029-907-7	Sequence 7, Appl
35	1583.4	87.8	8638	13	US-10-029-907-24	Sequence 24, Appl
36	1583.4	87.8	8638	13	US-10-029-907-25	Sequence 25, Appl
37	1583.4	87.8	8638	15	US-10-309-561-6	Sequence 6, Appl
38	1583.4	87.8	8638	15	US-10-309-561-7	Sequence 7, Appl
39	1583.4	87.8	8638	15	US-10-309-561-24	Sequence 24, Appl
40	1583.4	87.8	8638	15	US-10-309-561-25	Sequence 25, Appl
41	1583.4	87.8	8638	19	US-10-789-355-6	Sequence 6, Appl
42	1583.4	87.8	8638	19	US-10-789-355-7	Sequence 7, Appl
43	1583.4	87.8	8638	19	US-10-789-355-24	Sequence 24, Appl
44	1583.4	87.8	8638	19	US-10-789-355-25	Sequence 25, Appl
45	1583.4	87.8	8638	20	US-10-686-835-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-10-712-479-1
; Sequence 1, Application US/10712479
; Publication No. US20040209283A1
; GENERAL INFORMATION:
; APPLICANT: Yagi, Yoshihiko
; APPLICANT: Sheets, Michael P.
; APPLICANT: Wells, Peter A.
; APPLICANT: Shelly, John A.
; APPLICANT: Poorman, Roger A.
; APPLICANT: Eggs, Dennis E.
; TITLE OF INVENTION: A Continuous-Read Assay for the Detection of De Novo HCV RNA Poly
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 02-333-A
; CURRENT APPLICATION NUMBER: US/10712.479
; CURRENT FILING DATE: 2003-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1803
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1803)
; OTHER INFORMATION:
US-10-712-479-1

Query Match 100.0%; Score 1803; DB 20; Length 1803;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGGCTAGCATGTCAATGTCTCTATACATGGACAGCGCCCTGATCACACCGTGGCTGGC	60
Db	1	ATGGCTAGCATGTCAATGTCTCTATACATGGACAGCGCCCTGATCACACCGTGGCTGGC	60
Qy	61	GAGGAAAGCAAGCTGCCCATCAACGGCTGAGCAACTCTTGTGCTGCCTCACCATAACTCG	120
Db	61	GAGGAAAGCAAGCTGCCCATCAACGGCTGAGCAACTCTTGTGCTGCCTCACCATAACTCG	120
Qy	121	GTCTATTCCAAACATCCCGCAGTGCAAAGCCTGCGGAGAGAAAGGTCACCTTTGACAGA	180
Db	121	GTCTATTCCAAACATCCCGCAGTGCAAAGCCTGCGGAGAGAAAGGTCACCTTTGACAGA	180
Qy	181	CTGCAAGTCTGGAGATCATACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGCTCC	240
Db	181	CTGCAAGTCTGGAGATCATACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGCTCC	240
Qy	241	ACAGTGAAGGCTAAACTGCTATCTGTAGAGCAAGCATCCAGCTCCAGCAGGCGCGTT	300
Db	241	ACAGTGAAGGCTAAACTGCTATCTGTAGAGCAAGCATCCAGCTCCAGCAGGCGCGTT	300
Qy	301	GCCAAATCCAAATTTGGCTATGGGGCAAGGACGCTCGGAGCCTATCCAGCAGGCGCGTT	360
Db	301	GCCAAATCCAAATTTGGCTATGGGGCAAGGACGCTCGGAGCCTATCCAGCAGGCGCGTT	360
Qy	361	AACCAATCCGCTCGGTGTGGAAGGACTTGTGGAGGACACTGACACACCAATTCAGACC	420
Db	361	AACCAATCCGCTCGGTGTGGAAGGACTTGTGGAGGACACTGACACACCAATTCAGACC	420
Qy	421	ACCATCATGGCAAAATAGAGTTTCTGCGTCCAAACGAGAAAGAGGCGCGCAACCA	480
Db	421	ACCATCATGGCAAAATAGAGTTTCTGCGTCCAAACGAGAAAGAGGCGCGCAACCA	480
Qy	481	GCTGCGCTCATGTTATCCAGACCTGGGAGTTCGTGTATCGGAGAGATGGCCCTTTAC	540
Db	481	GCTGCGCTCATGTTATCCAGACCTGGGAGTTCGTGTATCGGAGAGATGGCCCTTTAC	540
Qy	541	GACGTGGTTTCCACTCTTCCCTCAGGCGGTGATGGGCTCCTCATACGGATTCCAACTCT	600
Db	541	GACGTGGTTTCCACTCTTCCCTCAGGCGGTGATGGGCTCCTCATACGGATTCCAACTCT	600
Qy	601	CCTAAGCAGCGGTCGAGTTCCTGTGTAATACCTGGAAAGCAAGAAATGCCCTATGGGC	660
Db	601	CCTAAGCAGCGGTCGAGTTCCTGTGTAATACCTGGAAAGCAAGAAATGCCCTATGGGC	660
Qy	661	TTCTCATATGACACCGCTGTTTGGACTCAACGGTCACTGAGATGACATCCGTGTTGAG	720
Db	661	TTCTCATATGACACCGCTGTTTGGACTCAACGGTCACTGAGATGACATCCGTGTTGAG	720
Qy	721	GAGTCAATTTACCAATGTTGTGACTTTGGCCCCCGAGCTAGACAGGCCATAAGTTCGCTC	780
Db	721	GAGTCAATTTACCAATGTTGTGACTTTGGCCCCCGAGCTAGACAGGCCATAAGTTCGCTC	780
Qy	781	ACAGACGGCTCTATGTGCGGGGTCCCATGACTAACTCCAAAGGCGCAGAACTGCGGCTAT	840
Db	781	ACAGACGGCTCTATGTGCGGGGTCCCATGACTAACTCCAAAGGCGCAGAACTGCGGCTAT	840
Qy	841	CGCCGGTCCGCGCAGCGGGTGTGTGACGACTAGCTGCGGTAATACCTCACATGCTAC	900
Db	841	CGCCGGTCCGCGCAGCGGGTGTGTGACGACTAGCTGCGGTAATACCTCACATGCTAC	900
Qy	901	TTGAAGGCGCTGACGCTGTGAGCTCCAGCTCCAGGACTGACAGTCTCGTGAAT	960
Db	901	TTGAAGGCGCTGACGCTGTGAGCTCCAGCTCCAGGACTGACAGTCTCGTGAAT	960
Qy	961	GGAGACGACTTGTGTTATCTGTGAAGCGCGGAAACCCAAAGAGAGCGCGGCAAGCCTA	1020
Db	961	GGAGACGACTTGTGTTATCTGTGAAGCGCGGAAACCCAAAGAGAGCGCGGCAAGCCTA	1020
Qy	1021	CGAGTCTTCACGAGGCTATGACTAGGTACTTGCCCCCCTCGGGACCCGCCCAACCG	1080
Db	1021	CGAGTCTTCACGAGGCTATGACTAGGTACTTGCCCCCCTCGGGACCCGCCCAACCG	1080

Qy	1081	GAATACGACTTGGAGCTGATAACATCGTGTTCCTCAATGTGTGTCGTCGCACACGATGCA	1140
Db	1081	GAATACGACTTGGAGCTGATAACATCGTGTTCCTCAATGTGTGTCGTCGCACACGATGCA	1140
Qy	1141	TCTGGCAAAAGGCTGTAACCTACCGGTGACCCACCGTCCCTTCGCGGGGCTGG	1200
Db	1141	TCTGGCAAAAGGCTGTAACCTACCGGTGACCCACCGTCCCTTCGCGGGGCTGG	1200
Qy	1201	TGGGAGACGCTAGGCACACACTCCAGTCAACTCCTGCTAGGCAACATCATGTATGCG	1260
Db	1201	TGGGAGACGCTAGGCACACACTCCAGTCAACTCCTGCTAGGCAACATCATGTATGCG	1260
Qy	1261	CCCACCTTGTGGCAAGGATGATTCGTGAGTCACTTCTTCTCCATCCTCTAGGCCAG	1320
Db	1261	CCCACCTTGTGGCAAGGATGATTCGTGAGTCACTTCTTCTCCATCCTCTAGGCCAG	1320
Qy	1321	GAGCAACTTGA AAAAGCCCTGGATTGTGAGATCTACGGGGCTTGTTCATTTAGGCCA	1380
Db	1321	GAGCAACTTGA AAAAGCCCTGGATTGTGAGATCTACGGGGCTTGTTCATTTAGGCCA	1380
Qy	1381	CTTGACCTACCTCAGATCAATTGAAACGACTCCATGCTTCTTCATCCTCTTAGGCCAG	1440
Db	1381	CTTGACCTACCTCAGATCAATTGAAACGACTCCATGCTTCTTCATCCTCTTAGGCCAG	1440
Qy	1441	TACTCTCCAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCACCC	1500
Db	1441	TACTCTCCAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCACCC	1500
Qy	1501	TTCCGAGTCTGGAGACATCGGGCCAGAAGTGTCCGGCTTAAAGTTACTGTCCAGGGGGGG	1560
Db	1501	TTCCGAGTCTGGAGACATCGGGCCAGAAGTGTCCGGCTTAAAGTTACTGTCCAGGGGGGG	1560
Qy	1561	AGGGCCGCCAATTTGTGECAGTAACCTCTTCAACTGGGCAAGTAAAGTTACTGTCCAGGCTTAAATC	1620
Db	1561	AGGGCCGCCAATTTGTGECAGTAACCTCTTCAACTGGGCAAGTAAAGTTACTGTCCAGGCTTAAATC	1620
Qy	1621	ACTCCAAATTCGGCTCGCTCCGGCTGGACTTGTCCGGCTGGTTCGTTGCTGGCTACAGC	1680
Db	1621	ACTCCAAATTCGGCTCGCTCCGGCTGGACTTGTCCGGCTGGTTCGTTGCTGGCTACAGC	1680
Qy	1681	GGGGGAGACATATCACAGCCTGTCTGCTGCGCCGACCCGCTGGTTCTATGTTGTCCTA	1740
Db	1681	GGGGGAGACATATCACAGCCTGTCTGCTGCGCCGACCCGCTGGTTCTATGTTGTCCTA	1740
Qy	1741	CTCCTACTCTCCGTAGGGGTAGGCATCTATCTACTCCCAACCGGCATCACCATCACCAT	1800
Db	1741	CTCCTACTCTCCGTAGGGGTAGGCATCTATCTACTCCCAACCGGCATCACCATCACCAT	1800
Qy	1801	CAC 1803	
Db	1801	CAC 1803	

RESULT 2
US-10-712-479-3
; Sequence 3, Application US/10712479
; Publication No. US20040209283A1
; GENERAL INFORMATION:
; APPLICANT: Yagi, Yoshihiko
; APPLICANT: Sheets, Michael P.
; APPLICANT: Wells, Peter A.
; APPLICANT: Shelly, John A.
; APPLICANT: Poorman, Roger A.
; APPLICANT: Epps, Dennis E.
; TITLE OF INVENTION: A Continuous-Read Assay for the Detection of De Novo HCV
; FILE REFERENCE: 02-333-A
; CURRENT APPLICATION NUMBER: US/10/712,479
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3

LENGTH: 1740
TYPE: DNA
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: C-terminally truncated HCV NS5B polymerase (C delta 21 NS5B)
NAME/KEY: CDS
LOCATION: (1)..(1740)
OTHER INFORMATION:
US-10-712-479-3

Query Match 95.5%; Score 1722; DB 20; Length 1740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTAGCATGTCATATGCTCTATACATGAGCAGGCGCCCTGATCACACCGTGGCTGCG 60
DB 1 ATGGCTAGCATGTCATATGCTCTATACATGAGCAGGCGCCCTGATCACACCGTGGCTGCG 60

QY 61 GAGGAAAGCAAGCTGCCATCAACCGCTGAGCAACTCTTCTGCTGCTCACCATAACCTG 120
DB 61 GAGGAAAGCAAGCTGCCATCAACCGCTGAGCAACTCTTCTGCTGCTCACCATAACCTG 120

QY 121 GTCTATTCCACAACTCCCGCAGTGCAGAGCTGCGGAGAGAGAGTCACTTTTGACAGA 180
DB 121 GTCTATTCCACAACTCCCGCAGTGCAGAGCTGCGGAGAGAGAGTCACTTTTGACAGA 180

QY 181 CTGCAAGTCTGGACGATCATTTACCGGACGCTGCTCAAGGAGATGAAGCGAAGCGCTCC 240
DB 181 CTGCAAGTCTGGACGATCATTTACCGGACGCTGCTCAAGGAGATGAAGCGAAGCGCTCC 240

QY 241 ACAGTGAAGCTTAACTGCTATCTGTAGAGAGCATGCAAGCTGACGCCGCCCGCATTCG 300
DB 241 ACAGTGAAGCTTAACTGCTATCTGTAGAGAGCATGCAAGCTGACGCCGCCCGCATTCG 300

QY 301 GCCAAATCCAAATTTGGCTATGGGCAAGAGCGTCCGAGCGCTATCAGAGCGGCGCTT 360
DB 301 GCCAAATCCAAATTTGGCTATGGGCAAGAGCGTCCGAGCGCTATCAGAGCGGCGCTT 360

QY 361 AACCATATCCGCTCGTGTGAAGGACTTGTGGAGGACACTGACACACCAATTCAGACC 420
DB 361 AACCATATCCGCTCGTGTGAAGGACTTGTGGAGGACACTGACACACCAATTCAGACC 420

QY 421 ACCATATGCGCAAAATAGGCTTTCTGCGTCCAAACAGAGAGAGGCGCGCAACCA 480
DB 421 ACCATATGCGCAAAATAGGCTTTCTGCGTCCAAACAGAGAGAGGCGCGCAACCA 480

QY 481 GCTGCGCTCATGCTATCCAGACCTGGAGTTCGTGTATCGAGAGATGGCCCTTAC 540
DB 481 GCTGCGCTCATGCTATCCAGACCTGGAGTTCGTGTATCGAGAGATGGCCCTTAC 540

QY 541 GAGTGGTTTCCACTTCTCAGGCGGTGATGGGCTCTCATAGGATTCATATCTCT 600
DB 541 GAGTGGTTTCCACTTCTCAGGCGGTGATGGGCTCTCATAGGATTCATATCTCT 600

QY 601 CCTAAGCAGCGGTGAGTTCCTGTGTAATACCTGGAAGCAAGAAATGCCCTATGGC 660
DB 601 CCTAAGCAGCGGTGAGTTCCTGTGTAATACCTGGAAGCAAGAAATGCCCTATGGC 660

QY 661 TTCTCATATGACACCGCTGTTTGTGACTCAAGGCTCACTGAGATGACATCCGTTGAG 720
DB 661 TTCTCATATGACACCGCTGTTTGTGACTCAAGGCTCACTGAGATGACATCCGTTGAG 720

QY 721 GAGTCAATTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCGCATAGTGGCTC 780
DB 721 GAGTCAATTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCGCATAGTGGCTC 780

QY 781 ACAGAGCGGCTCTATGTCGGGGTCCCATGACTCACTCCAAAGGCGAGACTGGGCTAT 840
DB 781 ACAGAGCGGCTCTATGTCGGGGTCCCATGACTCACTCCAAAGGCGAGACTGGGCTAT 840

QY 841 CGCCGGTCCCGGAGCGGCTGCTACAGCTAGCTGCGGTAATACCTCACAATGCTAC 900
DB 841 CGCCGGTCCCGGAGCGGCTGCTACAGCTAGCTGCGGTAATACCTCACAATGCTAC 900

RESULT 3

US-09-838-386-11

; Sequence 11, Application US/09838386

; Patent No. US20010055756A1

; GENERAL INFORMATION:

; APPLICANT: Kuko, George

; APPLICANT: Kuko, George

; TITLE OF INVENTION: Internal

; TITLE OF INVENTION: Internal

; FILE REFERENCE: 1011.218001

; CURRENT APPLICATION NUMBER: US/09/838,386

; CURRENT FILING DATE: 2001-04-20

; PRIOR APPLICATION NUMBER: US 60/198,793

; PRIOR FILING DATE: 2000-04-21

; Sequence 11, Application US/09838386

; Patent No. US20010055756A1

; GENERAL INFORMATION:

; APPLICANT: Kuko, George

; APPLICANT: Kuko, George

; TITLE OF INVENTION: Internal

; TITLE OF INVENTION: Internal

; FILE REFERENCE: 1011.218001

; CURRENT APPLICATION NUMBER: US/09/838,386

; CURRENT FILING DATE: 2001-04-20

; PRIOR APPLICATION NUMBER: US 60/198,793

; PRIOR FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: HTaA5B polymerase
; NAME/KEY: CDS
; LOCATION: (1)..(1881)
US-09-838-386-11

Query Match 89.2%; Score 1607.4; DB 9; Length 1884;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 8 GCATGTCAATGCTCTATACATGGACAGCGCCCTGATCACAACCGTGGCTGGCGAGGAAA 67
DB 104 GCTGCTCGATGCTCTACATGGACAGCGCCCTGATCAGCCATGGCGCGAGGAAA 163

QY 68 GCAAGTCCCATCAACCGCTGAGCACTCTTCTGCTGGTCAACATAAAGCTGCTATT 127
DB 164 GCCAGCTGCCCATCAACCGCTGAGCAACTCTTGGTGGTCAATCGCAACATGGTCTATT 223

QY 128 CCACAACATCCCGAGTGCAGAGCTGGCGGAGAGAGAGTCAACCTTTGACAGACTGCAAG 187
DB 224 CCACAACATCCCGAGCGCGCCCTGCGGAGAGAGAGTCACTTTGACAGACTGCAAG 283

QY 188 TCCTGGACGATCAITACCGGGACGTGCTCAAGGAGATGAAGCGGAGAGCGGTCACAGTGA 247
DB 284 TCCTGGACGACCACTACCGGGACGTGCTCAAGGAGATGAAGCGGAGAGCGGTCACAGTGA 343

QY 248 AGGCTAAACTGCTATCTGTAGAAGAGCATGCAAGTCAAGTCAAGCGCCCGCGATTGGCGCAAT 307
DB 344 AGGCCAAACTACTATCAGTAGAAGAGCGCTTAAGCTGACGCGCCCGCACATTCGCGCCAAAT 403

QY 308 CCAAAATTTGGCTATGGGCGAAGGAGCTCGGAGCGCTATCCAGAGGCGCGTTAACCAACA 367
DB 404 CCAAGTTTGGCTATGGGCGAAGGAGCGTCCGGAACCTATCCAGAGAGCGCGTTGACCAACA 463

QY 368 TCCGCTCCGCTGGGAAGACTTGTCTGGAGGACACTGACACACCAATTCAGACCAACCATCA 427
DB 464 TCCGCTCCGCTGGGAAGACTTGTCTGGAGGACACTGAGAGACACTGAGACCAACCATCA 523

QY 428 TGGCAAAAATAGAGTTTCTGCTCCAAACAGAGAAAGGAGCGGCGAAACAGAGTCCGC 487
DB 524 TGGCGAAAATAGAGTTTCTGCTCCAAACAGAGAAAGGAGCGGCGAAACAGAGTCCGC 583

QY 488 TCATCGTATTCAGACACTGGAGTTGCTGTATGGAGAGAGTGGCCCTTTACGACGTGG 547
DB 584 TTATCGTATTCAGACACTGGAGTTGCTGTATGGAGAGAGTGGCCCTTTACGACGTGG 643

QY 548 TTTCACACTCTTCTCAGCGCGTGATGGGCTCTCTATACGGATTCCTCAATACTCTCTAAGC 607
DB 644 TCTCCACCTTCTCAGCGCGTGATGGGCTCTCTATACGGATTCCTCAATACTCTCTAAGC 703

QY 608 AGCGGTTCAGTTCTCTGTTGAATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCAT 667
DB 704 AGCGGTTCAGTTCTCTGTTGAATACCTGGAATCAAGAAATGCCCTATGGGCTTCTCAT 763

QY 668 ATGACACCCGCTGTTTGTGACTCAACGGTCACTGAGAATGACATCCGTTGTGAGGAGTCAA 727
DB 764 ATGACACCCGCTGTTTGTGACTCAACGGTCACTGAGAGGACATCCGTTGTGAGGAGTCAA 823

QY 728 TTTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCCATAGGCTCGCTCAGAGGC 787
DB 824 TCTACCAATGTTGTGACTTGGCCCCCGAAGCCAGCAGACAGGCTATAAAGTGGCTCAGAGGC 883

QY 788 GGCTCTATGTCGGGGTCCCATGACTTAACCTCCAAAGGGCAGAACTCGCGCTATCCCGGT 847
DB 884 GGCTCTATATCGGGGGTCCCTTGACCAATTTCAAAAGGGCAGAACTCGCGCTATCCCGGT 943

QY 848 GCCGCGAGCGCGTGTGACGACTAGTGTGGGTAAATACCTTCATCTGCTACTTGAAGG 907
DB 944 GCCGCGAGCGCGTGTGCTGACGACGAGCTGGGTAAATACCTTCATCTGCTACTTGAAGG 1003

QY 908 CGGCTGCAGCCTGTGAGCTGGCAAGCTCCAGGACTGCAACGATGCTGTAATGGAGACG 967
DB 1004 CCTCTGCGGCTGTGAGCTGCCAGCTCCAGGACTGCAACGATGCTGTAACGGAGACG 1063

QY 968 ACCTTGTGTTATCTGTGTAAGCGGGGAAACCAAGAGACGCGGAGACCTACAGTCT 1027
DB 1064 ACCTTGTGTTATCTGCGAGAGCGGGAAACCAAGAGGATGCGGCAACCTACGAGTCT 1123

QY 1028 TCACGAGGCTATGACTAGTACTCTGCCCCCTGGGGACCGCCCCCAACCGGAATACG 1087
DB 1124 TCACGAGGCTATGACTAGTACTCTGCCCCCTGGGGACCTGCCCAACCGAATACG 1183

QY 1086 ACTTGGAGCTGATAACATCGTGTCTTCCAATGTGTGCGTTCGCACACGATGCATCTGGCA 1147
DB 1184 ACTTGGAGTTGATAACATCATGCTCTCTCAATGTGTGCGTTCGGGACGATGCATCGGCA 1243

QY 1148 AAGGGTGTACTACTACCTCACCCGTGACCCCAACCGTCCCCCTTGGCGGGCTGGGTGGAGA 1207
DB 1244 AAGGGTGTACTACTACCTCACCCGTGACCCCAACCGTCCCCCTTGGCGGGCTGGGTGGAGA 1303

QY 1208 CAGCTAGGCACACTCCAGTCACTCTGCTGCTAGGCAACATCATCATGTATGCGCCACCT 1267
DB 1304 CAGCTAGACACACTCCAATCAACTCTCTGCTAGGCAATATCATCATGTATGCGCCACCT 1363

QY 1268 TGTGGGCAAGGATGATTCTGATGACTCACTTCTTCTCCATCTTCTAGCCCCAGAGCAAC 1327
DB 1364 TATGGGCAAGGATGATTCTGATGACTCACTTCTTCTCCATCTTCTAGCCCCAGAGCAAC 1423

QY 1328 TTGAAAAGCCCTGGATGTGAGATCTAGGGGCTTGTACTCCTATGAGGCCACTTTGACC 1387
DB 1424 TTGAAAAGCCCTGGATGTGAGATCTAGGGGCTTGTACTCCTATGAGGCCACTTTGACC 1483

QY 1388 TACCTCAGATCATTTGAACGACTCCATGGTCTTAGCGCATTTTCACTCCATGATTACTCTC 1447
DB 1484 TACCTCAGATCATTTGAACGACTCCATGGTCTTAGCGCATTTTCACTCCATGATTACTCTC 1543

QY 1448 CAGGTGAGATCAATAGGTTGGCTTCATGCTCAGGAAGCTTGGGGTACCAACCTTGGGAG 1507
DB 1544 CAGGTGAAATCAATAGGTTGGCTTCATGCTCAGGAAGCTTGGGGTACCAACCTTGGGAG 1603

QY 1508 TCTGGACACATCGGCGCCAGAAAGTTCGCGCTAAAGTTTACTCTCCAGGGGGGAGGGCGG 1567
DB 1604 TCTGGAGACATCGGCGCCAGAAAGTTCGCGCTAAAGTTTACTCTCCAGGGGGGAGGGCTG 1663

QY 1568 CCATTTTGTGGCAAGTACTCTTCAAATGGGCACTAAGGACCAAGCTTAAACTCACTCCAA 1627
DB 1664 CCATTTGTGCAAGTACTCTTCAAATGGGCACTAAGGACCAAGCTTAAACTCACTCCAA 1723

QY 1628 TTCGGCTCGCTCCCGCTCGACTTGTTCGGCTGTTGTTGCTGCTGAGCTACAGCGGGGAG 1687
DB 1724 TCCCGCTCGCTCCCGCTCGACTTGTTCGGCTGTTGTTGCTGCTGAGCTACAGCGGGGAG 1783

QY 1688 ACATATATCAGGCTGTCTCGTCCCGACCGCGCTGGTTTCATGTTGTGCTACTCTCTAC 1747
DB 1784 ACATATATCAGGCTGTCTCGTCCCGACCGCGCTGGTTTCATGTTGTGCTACTCTCTAC 1843

QY 1748 TCTCCGTAGGGTAGGCATCTATCTACTCCCCCAACCG 1784
DB 1844 TTTCTGAGGGTAGGCATTTTACTGCTCCCCCAACCG 1880

RESULT 4
US-10-492-178-5
; Sequence 5, Application US/10492178
; Publication No. US20040247615A1
; GENERAL INFORMATION:
; APPLICANT: Emili A.
; APPLICANT: Kaslow, David C.
; APPLICANT: Bett, Andrew J.

```
/ APPLICANT: Shiver, John W.
/ APPLICANT: Nicotia, Alfredo
/ APPLICANT: Lahm, Armin
/ APPLICANT: Luzzago, Alessandra
/ APPLICANT: Cortese, Riccardo
/ APPLICANT: Colloca, Stefano
/ TITLE OF INVENTION: HEPATITIS C VIRUS VACCINE
/ FILE REFERENCE: ITR0015VP
/ CURRENT APPLICATION NUMBER: US/10/492,178
/ CURRENT FILING DATE: 2004-04-07
/ PRIOR APPLICATION NUMBER: PCT/US02/32512
/ PRIOR FILING DATE: 2002-10-10
/ PRIOR APPLICATION NUMBER: 60/363,774
/ PRIOR FILING DATE: 2002-03-13
/ PRIOR APPLICATION NUMBER: 60/328,655
/ PRIOR FILING DATE: 2001-10-11
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 5955
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: NS cDNA sequence
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(5955)
US-10-492-178-5

Query Match      88.6%; Score 1597.8; DB 20; Length 5955;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1665; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy      8 GCATGTCCTAATGTCCTATACATGAGCAGCGCGCTGATCACACCGTGCCTCGGAGGAAA 67
Db      4178 GCTGCTCAATGTCCTACACATGAGCAGCGCGCTTGTATCAGCCATGCTCGGAGGAAA 4237

Qy      68 GCAAGTCGCCATCAACGCGTGAACAACTCTCTGCTCGGTACCATAACTGCTCTATT 127
Db      4238 GCAAGTCGCCATCAACGCGTGAACAACTCTCTGCTCGGTACCATAACTGCTCTATT 4297

Qy      128 CCACAAATCCCGAGTCAAGCCCTGCGGAGAGAGAGTCAACCTTTGACAGCTGCAAG 187
Db      4298 CCACAAATCCCGAGTCAAGCCCTGCGGAGAGAGAGTCAACCTTTGACAGCTGCAAG 4357

Qy      188 TCCTGGAGCATATTACCGGAGCTGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTGA 247
Db      4358 TCCTGGAGCATATTACCGGAGCTGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTGA 4417

Qy      248 AGGCTAAACTGCTATCTGTAGAAGCATGCAAGCTGACGCGCCCGCATTCGGCCAAAT 307
Db      4418 AGGCTAAACTGCTATCTGTAGAAGCATGCAAGCTGACGCGCCCGCATTCGGCCAAAT 4477

Qy      308 CCAAATTTGGCTATGGGCGAAAGACGCTCGGAGAGCTTATCCAGCAGGCGCGTTAAACACA 367
Db      4478 CCAAATTTGGCTATGGGCGAAAGACGCTCGGAGAGCTTATCCAGCAGGCGCGTTAAACACA 4537

Qy      368 TCCGCTCCGTGTGAAGGACTTGTGTGAGGACATGACACACAAATTCAGACCAACATCA 427
Db      4538 TCCACTCCGTGTGAAGGACTTGTGTGAGGACATGACACACAAATTCAGACCAACATCA 4597

Qy      428 TGGCAAAAATAGAGTTTCTCGGTCCAAACAGAGAAAGAGCGCGCAACACAGCTCGCC 487
Db      4598 TGGCAAAAATAGAGTTTCTCGGTCCAAACAGAGAAAGAGCGCGCTAAGCGACGCGCCG 4657

Qy      488 TCATCGTATTCAGAGCTGGAGTTGCTGTATGCGAGAGATGGCGCTTTAGCAGGTGG 547
Db      4658 TTATCGTATTCAGAGCTGGAGTTGCTGTATGCGAGAGATGGCGCTTTATGATGG 4717

Qy      548 TTTCCACTCTTCTCAGGCGCTGATGGGCTCTCTATACAGGATTCCTATCTCTTAAC 607
Db      4718 TCTCCACCTCTCTCAGGCTGCTGATGGGCTCTCTATACAGGATTCCTATCTCTTGGGC 4777

Qy      608 AGCGGTCGAGTTCCTGGTGAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 667
Db      4778 AGCGGTCGAGTTCCTGGTGAATACCTGGAAATCAAGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 4837

Qy      668 ATGACACCGCGTGTGTTGACTCAACGCGTCACTGAGAAATGACATCCGTGTGAGGAGTCAA 727
Db      4838 ATGACACTCGGTGTTTTCGACTCAACGCTCACCGAGAACGACATCCGTGTGAGGAGTCAA 4897

Qy      728 TTTTACCAATGTTGTCAGTTCGCCCCCGAAGCTAGACAGGCCAATAAGGTGCTCACAGAGC 787
Db      4898 TTTTACCAATGTTGTCAGTTCGCCCCCGAAGCTAGACAGGCCAATAAAATCGCTCACAGAGC 4957

Qy      788 GCGCTCTATGTCGGGCGTCCCATGACTAATCTCAAGAGCGAGAACTGCGGCTATCCCGGT 847
Db      4958 GCGCTTATATCGGGGCTCTCTGACTAATCAAAAGGGCAGAACTGCGGCTATCCCGGT 5017

Qy      848 GCGCGAGCGGGGCTGCTGACGACTAGCTGCGGTAAATACCTCACAATGCTACTTGAAGG 907
Db      5018 GCGCGAGCGGGGCTGCTGACGACTAGCTGCGGTAAACACCTCACAATGCTACTTGAAGG 5077

Qy      908 CCGCTGACGCTGTCGAGCTGCCAAGTCCAGGACTGCACGATGCTCGTGAATGGAGACG 967
Db      5078 CCGCTGACGCTGTCGAGCTGCCAAGTCCAGGACTGCACGATGCTCGTGAATGGAGACG 5137

Qy      968 ACCTTGTGCTTATCTGTGAAAGCGCGGAAACCCAAAGAGGACGCGGAGCCTACGAGTCT 1027
Db      5138 ACCTTGTGCTTATCTGTGAAAGCGCGGAAACCCAAAGAGGACGCGGAGCCTACGAGTCT 5197

Qy      1028 TCACGAGGCTATGACTAGGTACTCTGCCCCCTCGGGAGCCCGCCCAACCGGAAATACG 1087
Db      5198 TCACGAGGCTATGACTAGGTACTCTGCCCCCTCGGGAGCCCGCCCAACCGGAAATACG 5257

Qy      1088 ACTTGGAGCTGATAAACAATCGTGTTCCTCAATGTCGCTGCGACACAGTGCATCTGGCA 1147
Db      5258 ACTTGGAGCTGATAAACAATCGTGTTCCTCAATGTCGCTGCGACACAGTGCATCTGGCA 5317

Qy      1148 AAAGGCTGTACTACTCTCACCCGCTGACCCACCGTCCCTTTCGCGGCGCTCGCTGGGAGA 1207
Db      5318 AAAGGCTGTACTACTCTCACCCGCTGACCCACCGTCCCTTTCGCGGCGCTCGCTGGGAGA 5377

Qy      1208 CAGCTAGGACACATCCAGTCAACTCTGCGGTAGGCAACATCATATGATGATGGGCCACTT 1267
Db      5378 CAGCTAGACACATCCAGTCAACTCTGCGGTAGGCAACATCATATGATGATGGGCCACTT 5437

Qy      1268 TGTGGGCAAGGATTTCTGATGACTCACTTCTTCTCCATCTTCTAGCCCACTTGACC 1387
Db      5438 TGTGGGCAAGGATTTCTGATGACTCACTTCTTCTCCATCTTCTAGCCCACTTGACC 5497

Qy      1328 TTGAAAAGCCCTGGATTTGTGAGTCTAGCGGCTTGTACTCCATTTAGCCCACTTGACC 1387
Db      5498 TTGAAAAGCCCTGGATTTGTGAGTCTAGCGGCTTGTACTCCATTTAGCCCACTTGACC 5557

Qy      1388 TACCTCAGATCATTTGAACGACTCCATGGTCTTAGCGCATTTTCACTCCATAGTACTCTC 1447
Db      5558 TACCTCAGATCATTTGAACGACTCCATGGTCTTAGCGCATTTTCACTCCATAGTACTCTC 5617

Qy      1448 CAGGTGAGATCAATAGGTTGGCTTATGCTCAGGAGCTTGGGGTACCACTTCCGAG 1507
Db      5618 CAGGTGAGATCAATAGGTTGGCTTATGCTCAGGAGCTTGGGGTACCACTTCCGAG 5677

Qy      1508 TCTGGAGACATCGGGCGAGAGTGTCCGCGCTAAGTACTGTCCAGGGGGGGAGGGCGG 1567
Db      5678 TCTGGAGACATCGGGCGAGAGCTGTCCGCGCTAAGTACTGTCCAGGGGGGGAGGGCGG 5737

Qy      1568 CCATTTGTGGCAAGTACCTCTTCAACTGGGCAAGTGAAGCAACCAAGCTTAACTCCTCCAA 1627
Db      5738 CCATTTGTGGCAAGTACCTCTTCAACTGGGCAAGTGAAGCAACCAAGCTTAACTCCTCCAA 5797

Qy      1628 TTTCCGCTGCTGCTCCGCGTGGACTTGTCCGCTGCTTGTCTGCTGCTACAGCGGGGAG 1687
Db      5798 TTTCCGCTGCTGCTCCGCGTGGACTTGTCCGCTGCTTGTCTGCTGCTACAGCGGGGAG 5857

Qy      1688 ACATATATACAGAGCTGCTCTGTCGCCGACCCCGCTGGTTCATGTTGTGCTCTACTCTAC 1747
```

```
Db 5858 ACATATATACAGCGTGTCTCGTGCCGAGACCCCGCTGGTTCATGCTGTGCTACTCTCTAC 5917
Qy 1748 TCTCGTAGGGGTAGGCATCTATCTACTTCTCCCAACCG 1784
Db 5918 TTTCTGTAGGGTAGGCATCTACTCTCTCCCAACCG 5954

RESULT 5
US-10-492-178-4
; Sequence 4, Application US/10492178
; Publication No. US20040247615A1
; GENERAL INFORMATION:
; APPLICANT: Emini, Emilio A.
; APPLICANT: Kaslow, David C.
; APPLICANT: Bett, Andrew J.
; APPLICANT: Shiver, John W.
; APPLICANT: Nicosia, Alfredo
; APPLICANT: Lahm, Armin
; APPLICANT: Luzzago, Alessandra
; APPLICANT: Cortese, Riccardo
; APPLICANT: Colloca, Stefano
; TITLE OF INVENTION: HEPATITIS C VIRUS VACCINE
; FILE REFERENCE: ITR0015YP
; CURRENT APPLICATION NUMBER: US/10/492,178
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: PCT/US02/32512
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/363,774
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/328,655
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 37090
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: MRKad6-NSmut nucleic acid
US-10-492-178-4

Query Match 88.3%; Score 1592.8; DB 20; Length 37090;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Qy 8 GCATGTCAATGTCTATACATGGACAGGCGCCCTGATCACACGGTGGTGGGAGGAAA 67
Db 5441 GCTGCTCAATGTCTTACATGGACAGGCGCCCTTGATCAGCCATCGCTGGGAGGAAA 5500
Qy 68 GCAAGCTGCCCATCAACGGCGCTGAGCAATCTCTTGTCTGGGTACCAATACCTGGTCTATT 127
Db 5501 GCAAGCTGCCCATCAACGGCGCTTGAGCAATCTCTTGTCTGGGCCCAACCAATACATGTTTATG 5560
Qy 128 CCACAACATCCCGAGTGGCAAGCTTGGGAGAGAGGTTCACCTTTTGACAGACTCCAAAG 187
Db 5561 CCACAACATCTCGCAGCGCAGGCGCTTGGGAGAGAGGTTCACCTTTTGACAGACTCCAAAG 5620
Qy 188 TCCTGGAGCATCATTAACGGGAGCTGTCTCAAGGAGATGAAGGCGAGGCGTCCACAGTGA 247
Db 5621 TCCTGGAGCAGCCACTACGGGAGCTGTCTCAAGGAGATGAAGGCGAGGCGTCCACAGTTA 5680
Qy 248 AGGCTAAACTGCTATCTGTAGAGAGCATGCAAGCTGACGCCCCCGCATTCGGCCAAAT 307
Db 5681 AGGCTAAACTCTTATCTGTAGAGAGCGCTGCAAGCTGACGCCCCCACTTCGGCCAAAT 5740
Qy 308 CCAAAATTTGGCTATGGGGCAAGGACGTCCGGAGCCCTTACGACAGGCGGTTTAAACACA 367
Db 5741 CCAAGTTTGGCTATGGGGCAAGGACGTCCGGAACTTATCCAGCAAGGCGGTTTAAACACA 5800
Qy 368 TCCGCTCCGTGTGAGGACTTGTCTGGAGGACCTGACACACCAATTCAGACCAACCATCA 427
Db 5801 TCCACTCCGTGTGAGGACTTGTCTGGAAGACCTGTGACACCAANTTGCACCAACCATCA 5860
```

```
Qy 428 TGGCAAAAAATAGGTTTTTCTGCGTCCAAACCCAGAGAAAGAGCGCCGCAAAACGAGCTCGCC 487
Db 5861 TGGCAAAAAATAGGTTTTTCTGTTGCCAACACAGAGAAAGAGCGCCGTAAGCCAGCGCCGCC 5920
Qy 488 TCATCGTATTTCCAGACCTCGGAGTTCTGTATGCGAGAAGATGGCCCTTTACGAGCTGG 547
Db 5921 TTATCGTATTTCCAGATCTGGGAGTCCGTGTATGCGAGAAGATGGCCCTCTATGATGTGG 5980
Qy 548 TTTCCACTCTTCTCAGGCGGTGATGGGCTCCCTCATACCGATTTCCAAATATCTCTCTAAGC 607
Db 5981 TCTCCACCCCTTCTCAGGCTCGTGTATGGGCTCCCTCATACCGATTTCCAGTATCTCTCTCGGC 6040
Qy 608 AGCGGGTCGAGTTTCTCTGTTGAATFACCTGGAAAGCAAAAGAAATGCCCCATGCGGCTTCTCAT 657
Db 6041 AGCGAGTCGAGTTTCTCTGTTGAATFACCTGGAAATCAAAAGAAACCCCATGGGCTTTTCAT 6100
Qy 668 ATGACACCCCGCTGTCTTTTGACTCAACCGTCACTGAGAAATGACATCCGTTGTGAGGAGTCAA 727
Db 6101 ATGACACTCGCTGTTTCGACTCAACGGTCAACCGAGAACGACATCCGTTGTGAGGAGTCAA 6160
Qy 728 TTTTACCAATGTTGTGACTTGGCCCCGGAAGCTAGACAGGCGCATAAAGTCCCTCACAGAGC 787
Db 6161 TTTTACCAATGTTGTGACTTGGCCCCGGAAGCCAGACAGGCGCATAAATCCCTCACAGAGC 6220
Qy 788 GGCTCTATGTCGGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGCGGCTATTCGCCGCT 847
Db 6221 GGCTTTATTCGGGGGTCTCTGACTAAATTCAAAGGGCGAGAACTGCGGCTATTCGCCGCT 6280
Qy 848 GCGCGGAGCGGGCTGCTGACGACTAGCTGCGGTATATACCCCTCACATGCTACTTGAAGG 907
Db 6281 GCCCGCGAGCGGGCTGCTGACGACTAGCTGCGGTATATACCCCTCACATGCTACTTGAAGG 6340
Qy 908 CCGCTGCAGCCTGTCGAGCTGCCAAGCTCCAGGATGTCACGATGCTCGTGAATGGAGAGC 967
Db 6341 CCTCTGCAGCCTGTCGAGCTGCCAAGCTCCAGGACTGACGATGCTCGTGAATGGAGAGC 6400
Qy 968 ACCTTGTGTTTATCTGTAAAGCGCGGAAACCCAAAGAGGAGCGGCGAAGCTACGAGTCT 1027
Db 6401 GCCTTGTGTTTATCTGTAAAGCGCGGAAACCCAAAGAGGAGCGGCGGAGCTACGAGTCT 6460
Qy 1028 TCACGAGGCTATGACTAGGTACTCTGCCCGCTGAGGAGCGCCGCCCAACCGGAATACG 1087
Db 6461 TCACGAGGCTATGACTAGGTACTCTGCCCGCTGAGGAGCGCCGCCCAACCGGAATACG 6520
Qy 1088 ACTTGGAGCTGATTAACATCGTGTTCCTTCCAAATGTTGCGTGCACACGATCATCTTGGCA 114
Db 6521 ACTTGGAGCTGATTAACATCAATGTTCTTCCAAATGTTGCGTGCACACGATCATCAGGCA 6580
Qy 1148 AAAGGGTGTACTACTCTACCCGTGACCCCAACCGTCCCTTTGCCGGGCTGGTGGGAAA 120
Db 6581 AAAGGGTGTACTACTCTACCCGTGATCCACACCGCTTCCGAGGAGCGGCTGGTGGGAAA 664
Qy 1208 CAGCTAGGAGACATCCAGTCAACTCTGCTAGGCAACATCATCATGATGATGCGCCACTT 126
Db 6641 CAGCTAGACACACTCCAGTTAACTCTGCTAGGCAACATCATCATGATGATGCGCCACTT 6700
Qy 1268 TGTGGGCAAGGATGATTCTGATGACTCACTTCTCTCCATCTCTTCTAGCCCAAGGAGAAC 132
Db 6701 TGTGGGCAAGGATGATTCTGATGACTCACTTCTCTCCATCTCTTCTAGCAGGAGAAC 676
Qy 1328 TTTGAAAAAGCCCTTGGATTTGTACAGATCTACGGGGCTTGTGTCTTCCATCTCTAGCCCAAGGAGAAC 138
Db 6761 TTTGAAAAAGCCCTTGGATTTGTACAGATCTACGGGGCTTGTGTCTTCCATCTCTAGCAGGAGAAC 682
Qy 1388 TACCTCAGATCAATGAACGACTCCATGTTGTAGGCGCATTTTCACTCCATAGTTACTCTC 144
Db 6821 TACCTCAGATCAATGAACGACTCCATGTTGTAGGCGCATTTTCACTCCATAGTTACTCTC 688
Qy 1448 CAGGTGAGATCAATAGGGTGGCTTCTATGCTCAGGAAGCTTGGGGTACCACTTTGCGAG 150
Db 6881 CAGGTGAGATCAATAGGGTGGCTTCTATGCTCAGGAAGCTTGGGGTACCACTTTGCGAG 694
```

```
QY 1508 TCTGGAGACATCGGGCCAGAAAGTGTCCGGCTTAAGTTACTGTCCAGGGGGGAGGGCCG 1567
Db 6941 TCTGGAGACATCGGGCCAGAGCGTCCGGCTAGGCTACTGTCCAGGGGGGAGGGCCG 7000
QY 1568 CCAATTGTGGCAGTACTCTTCAACTGGGCGTAGGACCAAGCTTAAACTCACTCCAA 1627
Db 7001 CCACTTGTGGCAGTACTCTTCAACTGGGCGTAGGACCAAACTCAAACCTCACTCCAA 7060
QY 1628 TTCCGGCTCGTCCCGCTGGACTTGTCCGGCTGGTTCGTGGCTACAGCGGGGAG 1687
Db 7061 TCCGGCTCGTCCCGCTGGACTTGTCCGGCTGGTTCGTGGCTACAGCGGGGAG 7120
QY 1688 ACATATATCACAGCTGTCTCGTGGCCGACCCCGCTGGTTCATGTGTGCTACTCCTAC 1747
Db 7121 ACATATATCACAGCTGTCTCGTGGCCGACCCCGCTGGTTCATGTGTGCTACTCCTAC 7180
QY 1748 TCTCCGTAGGGTAGGCATCTATCTACTCCCAACCGGA 1787
Db 7181 TTTCTGTAGGGTAGGCATCTACTCTCCCAACCGGTA 7220

RESULT 6
US-10-492-178-2
; Sequence 2, Application US/10492178
; Publication NO. US20040247615A1
; GENERAL INFORMATION:
; APPLICANT: Emini, Emilio A.
; APPLICANT: Kaslow, David C.
; APPLICANT: Bett, Andrew J.
; APPLICANT: Shiver, John W.
; APPLICANT: Nicotia, Alfredo
; APPLICANT: Lahm, Armin
; APPLICANT: Luzzago, Alessandra
; APPLICANT: Cortese, Riccardo
; APPLICANT: Colloca, Stefano
; TITLE OF INVENTION: HEPATITIS C VIRUS VACCINE
; FILE REFERENCE: ITR0015VP
; CURRENT APPLICATION NUMBER: US/10/492,178
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: PCT/US02/32512
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/363,774
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/328,655
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 5965
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Non-optimized cDNA sequence encoding SEQ. ID. NO.
US-10-492-178-2

Query Match 88.3%; Score 1591.4; DB 20; Length 5965;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 8 GCATGTCAATGTCTTATACATGGAACGGCCCTGATCACACCGTGGCGTGGAGGAAA 67
Db 4184 GCTGCTCAATGTCTTATACATGGAACGGCCCTGATCACACCGTGGCGTGGAGAAA 4243
QY 68 GCAGCTGCCCATCAACGGCTGAGCAATCTTGTGCTGCTACCATAACTGGTCTATT 127
Db 4244 GCAAGCTGCCCATCAACGGCTGAGCAATCTTGTGCTGCTACCATAACTGGTCTATT 4303
QY 128 CCACAACATCCGCGAGTGAAGCTCGGCGAGAGAGAGGTCACTTTGACAGACTGCAAG 187
Db 4304 CCACAACATCCGCGAGTGAAGCTCGGCGAGAGAGAGGTCACTTTGACAGACTGCAAG 4363
QY 188 TCCTGGACGATCATTTACCGGACGCTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTGA 247
```

```
Db 4364 TCCTGGACGATCATTTACCGGACGCTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTTA 4423
QY 248 AGGCTAAACTCTCTATCTGTAGAAAGCATGCAAGCTCAGCGCCCGCATTTCCGGCAAAAT 307
Db 4424 AGGCTAAACTCTCTATCTGTAGAAAGCATGCAAGCTCAGCGCCCGCATTTCCGGCAAAAT 4483
QY 308 CCAAAATTTGGCTATCGGGCCAAAGGACGTCCGGAGCCTATCCAGAGGGCCGTTAAACCA 367
Db 4484 CCAAAATTTGGCTATCGGGCCAAAGGACGTCCGGAGCCTATCCAGAGGGCCGTTAAACCA 4543
QY 368 TCCGCTCGCTGTGGAGGACCTTGTGGAGGACCTGACACACCAATTCAGACCAACCATCA 427
Db 4544 TCCGCTCGCTGTGGAGGACCTTGTGGAGGACCTGACACCAATTCAGACCAACCATCA 4603
QY 428 TGGCAAAAATGAGGCTTTTCTGCGTCCAAACAGAGAAAGGAGGCGGCAAAACAGCTCGCC 487
Db 4604 TGGCAAAAATGAGGCTTTTCTGCGTCCAAACAGAGAAAGGAGGCGGTAAGCCAGCCCGCC 4663
QY 488 TCATCGTATTCACAGACCTGGGAGTTCGTGTATCGGAGAGATGGCCCTTTACGACGTGG 547
Db 4664 TTATCGTATTCACAGATCTGGGAGTTCGTGTATCGGAGAGATGGCCCTTCTATGATGG 4723
QY 548 TTTCCACTCTTCTCAGGCGGTGATGGCTCTCATACGGATTCGAATCTCTCTCTTAAGC 607
Db 4724 TCTCCACCTTCTCAGGCGGTGATGGCTCTCATACGGATTCGAATCTCTCTCTTGGC 4783
QY 608 AGCGGCTCGAGTCTCTGCTGAATACCTGGAAAGCAAGAAATGCGCTTATGGCTTCTCAT 667
Db 4784 AGCGAGTCTGAGTCTCTGCTGAATACCTGGAAATCAAAAGAAACCCATGGGCTTTTCT 4843
QY 668 ATGACACCCGCTGTTTGTGACTCAACGGTCACTGAGAATGACATCCGTGTTGAGGAGTCAA 727
Db 4844 ATGACACTCGCTGTTTGTGACTCAACGGTCACTGAGAATGACATCCGTGTTGAGGAGTCAA 4903
QY 728 TTTACCAATGTGTGACTTGGCCCGGAGCTAGAGGCGCATTAAGTTCGCTCACAGAGC 787
Db 4904 TTTACCAATGTGTGACTTGGCCCGGAGCTAGAGGCGCATTAAGTTCGCTCACAGAGC 4963
QY 788 GGCTCTATGTGGGGGTCCCATGACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 847
Db 4964 GGCTCTATGTGGGGGTCCCATGACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 5023
QY 848 GCGCGGAGCGGCTGCTGACGACTAGCTGCGGTATACCTCACTCACTCACTCACTCACTGAGG 907
Db 5024 GCGCGGAGCGGCTGCTGACGACTAGCTGCGGTATACCTCACTCACTCACTCACTTGAAGG 5083
QY 908 CCGCTCGAGCTGTTCGAGCTGCGCAAGCTCCAGGACTGCAAGTGTCTGCTCAATGGAGACG 967
Db 5084 CCGCTCGAGCTGTTCGAGCTGCGCAAGCTCCAGGACTGCAAGTGTCTGCTCAATGGAGACG 5143
QY 968 ACCTTGTGTTATCTGTGAAGCGGGGAAACCAAGAGGACGCGGCAAGCTTACGAGTCT 1027
Db 5144 GCCTTGTGTTATCTGTGAAGCGGGGAAACCAAGAGGACGCGGCAAGCTTACGAGTCT 5203
QY 1028 TCAGGAGGCTATGACTAGTACTTGTGCGGCTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 1087
Db 5204 TCAGGAGGCTATGACTAGTACTTGTGCGGCTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 5263
QY 1088 ACTTGAGCTGATAACATCGTGTCTCCAAATGCTGCTGCGGCTGCGGAGACGATCTGGCA 1147
Db 5264 ACTTGAGCTGATAACATCGTGTCTCCAAATGCTGCTGCGGCTGCGGAGACGATCTGGCA 5323
QY 1148 AAAGGCTGATACCTCACCGCTGACCCCAACCGTCCCGCTTGGCGGGCTGCGGAGGAGA 1207
Db 5324 AAAGGCTGATACCTCACCGCTGACCCCAACCGTCCCGCTTGGCGGGCTGCGGAGGAGA 5383
QY 1208 CAGCTAGGACACTCCAGTCAACTCTCTGCTAGGCAACATCATCATATGATGGCCCACTT 1267
Db 5384 CAGCTAGGACACTCCAGTCAACTCTCTGCTAGGCAACATCATCATATGATGGCCCACTT 5443
QY 1268 TGTGGCAAGGATGATCTGATGACTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1327
```

Db 5444 TGTGGGCMAGGATGATTTCTGATGACTCACTTCTTCTCCATCTCTTAGCACAGGACAA 5503
Qy 1328 TTGAAAAGCCCTGGATTGTGTCAGATCTACGGGGCTTGTTACTCCATTTAGCCACTTGACC 1387
Db 5504 TTGAAAAGCCCTGGACTGCCAGATCTACGGGGCTTGTTACTCCATTTAGCCACTTGACC 5563
Qy 1388 TACCTCAGATCATTGAACGACTCCATGCTCTTACGGGATTTTCACTCCTAGTTACTCTC 1447
Db 5564 TACCTCAGATCATTGAACGACTCCATGCTCTTACGGGATTTTCACTCCTAGTTACTCTC 5623
Qy 1448 CAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCACTTGGGAG 1507
Db 5624 CAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCACTTGGGAG 5683
Qy 1508 TCTGAGACATCGGGCCAGAAAGTGTCCGCGCTTAGTTACTTCTCCAGGGGGAGGGCCG 1567
Db 5684 TCTGAGACATCGGGCCAGAGAGCGTCCGCGTCTAGCTACTGTCAGGGGGAGGGCCG 5743
Qy 1568 CCATTTGTGGCAAGTACCTCTTCAACTGGGCAAGTAAAGCAAGCTTAAACTCACTCAA 1627
Db 5744 CCACTTGTGCAAGTACCTCTTCAACTGGGCAAGTAAAGCAAGCTTAAACTCACTCAA 5803
Qy 1628 TTCGGGCTGCTCCGGCTGGACTTGTCCGGTGGTTGGTTGCTGGCTACAGCGGGGAG 1687
Db 5804 TCCCGGCTGGCTCCAGCTGGACTTGTCCGGTGGTTGGTTGCTGGTTACAGCGGGGAG 5863
Qy 1688 ACATATATACAGAGCTGTCTGTGCGCGAGCCCGCTGGTTCTATGTTGCTACTTCTTAC 1747
Db 5864 ACATATATACAGAGCTGTCTGTGCGCGAGCCCGCTGGTTCTATGTTGCTACTTCTTAC 5923
Qy 1748 TCTCGTAGGGGTAGGCATCTATCTACTCTCCCAACCG 1784
Db 5924 TTCTGTAGGGGTAGGCATCTACTTGTCTCCCAACCG 5960

RESULT 7

US-10-259-275-41
; Sequence 41, Application US/10259275
; Publication No. US20030125541A1
; GENERAL INFORMATION:
; APPLICANT: Lemon, Stanley M.
; APPLICANT: Yi, Minkyung
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007 0120
; CURRENT APPLICATION NUMBER: US/10/259,275
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US 60/171,909
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/747,419
; PRIOR FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/325,236
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/338,123
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 6189
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding hepatitis C virus polypeptide
; OTHER INFORMATION: derived from Con1

US-10-259-275-41

Query Match 87.9%; Score 1585; DB 15; Length 6189;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
Qy 8 GCATGTCAATGTCTTACATGACAGGCGCCCTGTATCAGGCATCGGCTGCGGAGAAA 67
Db 4178 GCTGCTCGATGTCTTACATGACAGGCGCCCTGTATCAGGCATCGGCTGCGGAGAAA 4237

Qy 68 GGAAGCTGCCCATCAACGGCTGAGCAACTCCTTGCTGCTCACCATAAACCCTGGTCTATT 127
Db 4238 CCAAGCTGCCCATCAATGACTGAGCAACTCTTTGCTCCGTCAACCAACTTTGGTCTATG 4297
Qy 128 CCACAAATCCCGCAGTCAAGCTTCGGCAGAGAAAGTGTACCTTTTGACAGACTGCAAG 187
Db 4298 CTACAAATCTCGCAGCAGAGCTTCGGCAGAGAAAGTGTACCTTTTGACAGACTGCAAG 4357
Qy 188 TCTTGACGATCATTTACCGGAGCTGCTCAAGGAGNTGAAGGGAGGGTCCACAGTGA 247
Db 4358 TCTTGACGACCACTTACCGGAGCTGCTCAAGGAGNTGAAGGGAGGGTCCACAGTGA 4417
Qy 248 AGGCTAACTGCTATCTGTAGAGAAAGCATGCAAGCTGACGCCCGCCGCTATTCGGCCAAAT 307
Db 4418 AGGCTAACTTCTATCCGTGGAGAAAGCTGTAAAGCTGACGCCCGCCGCTATTCGGCCAGAT 4477
Qy 308 CCAATTTGGCTATCGGGCAAGCAGCTTCGGAGCCCTATCCAGCAGGGCGGTTTAAACACA 367
Db 4478 CTAAATTTGGCTATCGGGCAAGGAGCTTCGGAGACCTATCCAGCAAGGCGGTTTAAACACA 4537
Qy 368 TCCGCTCCGTGTGGAGGACTTCTCTGGAGGACACTGACACACCAATTCAGACACCATCA 427
Db 4538 TCCGCTCCGTGTGGAGGACTTCTCTGAGACACTGAGACACCAATTCAGACACCATCA 4597
Qy 428 TGGCAAAAATGAGGTTTCTTGGCTCAACAGAGAAAGGAGCGCCCAACACAGCTCGCC 487
Db 4598 TGGCAAAAATGAGGTTTCTTGGCTCAACAGAGAAAGGAGCGCGCCCAACACAGCTCGCC 4857
Qy 488 TCATCGTATTTCCAGAGCTCGGAGTTCTGTATGCGAGAAAGTGGCCCTTTACGAGGTGG 547
Db 4658 TTATCGTATTTCCAGAGTTTGGGGTTTGGTGTGCGAGAAATTTGCCCCCTTTACGATGTG 4717
Qy 548 TTTCCACTTCTCTCAGCGCGTGTGAGGCTCTCATACGATTTCCAAATCTCTCTAAAGC 607
Db 4718 TCTCCACCTTCCCTCAGCGCGTGTGAGGCTCTTATACGATTTCCAAATCTCTCTCGAC 4777
Qy 608 AGCGGTGCGATTCTCTGTGAATACCTGGAGAAAGCAAGAAATGCCCTATCGGCTTCTCAT 667
Db 4778 AGCGGTGCGATTCTCTGTGAATGCTTGGAAAGCGAAAGAAATGCCCTATCGGCTTCTCAT 4837
Qy 668 ATGACACCGCTGTGTTGACTCAACGGTCACTCAGAAATGACATCCGTGTTGAGGAGTCAA 727
Db 4838 ATGACACCGCTGTGTTGACTCAACGGTCACTCAGAAATGACATCCGTGTTGAGGAGTCAA 4897
Qy 728 TTTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCGCATTAAGTCTGCTACAGAGC 787
Db 4898 TCTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCGCATTAAGTCTGCTACAGAGC 4957
Qy 788 GGCTCTATGTCCGGGGTCCCATGACTAACTCCAAAGGGCGAGAACTGCGGCTATCGCCGGT 847
Db 4958 GGCTTTACATCGGGGGCCCCCTGACTAAATTTAAAGGGCGAGAACTGCGGCTATCGCCGGT 5017
Qy 848 GCCGCGAGCGCGTGTGACGACTGCTGCGTAAATACCTCAGATGCTACTTTGAGG 907
Db 5018 GCCGCGAGCGCGTGTGACGACTGCTGCGTAAATACCTCAGATGCTACTTTGAGG 5077
Qy 908 CCGCTGACCGCTGTGAGCTGCCAAGCTGCCAGGACTGCACCATGCTCTGTGAATGGAGACG 967
Db 5078 CCGCTGCGGCTGTGAGCTGCGAAGCTTCCAGGACTGCACCATGCTCTGTGAATGGAGACG 5137
Qy 968 ACCTTGTCTTATCTGTGAAAGCGCGGAACCCCAAGAGGACCGCGCAAGCCCTACGAGTCT 1027
Db 5138 ACCTTGTCTTATCTGTGAAAGCGCGGGGACCCCAAGAGGACGAGCGGAGCCCTACGGGCT 5197
Qy 1028 TCACGAGCGCTATGACTAGGTACTCTGCCCCCTTGGGGATCCCGCCCCAACCGGAATACG 1087
Db 5198 TCACGAGCGCTATGACTAGGTACTCTGCCCCCTTGGGGATCCCGCCCCAACCGGAATACG 5257
Qy 1088 ACTTGGAGCTGATTAACATCGTGTCTTCAATGTGCTCGGTGCGACACCATCATCTCGCA 1147
Db 5258 ACTTGGAGTTGATTAACATCATGCTCTTCAATGTGCTAGTGGCGCAGCATGCTATCTGGCA 5317
Qy 1148 AAAGGGTGTACTACTCTCACCGGTGACCCCGCTCCCGCTTGGCGGGCTGTGTGGGAGA 1207

Db 5138 ACCTTGCTGTTATCTGTGAAGCGGGGAGCCCAAGAGGACGAGCGGCGCTACGCGCCT 5197
Qy 1028 TCACGAGGCTATGACTAGGTACTCTGCCGCCCTGGGGACCCGCCCAACCGGAATACG 1087
Db 5198 TCACGAGGCTATGACTAGTACTCTGCCGCCCTGGGGACCCGCCCAACCGGAATACG 5257
Qy 1088 ACTTGGAGCTGATAACATCGTGTCTCTCAATGTGTGGTGGACAGATGCATCTGGCA 1147
Db 5258 ACTTGGAGTTGATAACATCATCTCTCAATGTGTGAGTGGCGCAGATGCATCTGGCA 5317
Qy 1148 AAAGGGTACTACTACCTACCGGTGACCCCAACCGTCCCTTGGCGGGCTGCGGGGAGA 1207
Db 5318 AAAGGGTACTATCTACCCGTGACCCCAACCGTCCCTTGGCGGGCTGCGTGGGAGA 5377
Qy 1208 CAGCTAGGCACACTCCAGTCAACTCTGGCTAGGCAACATCATATGTATGCGCCCACTT 1267
Db 5378 CAGCTAGACACACTCCAGTCAATCTGGCTAGGCAACATCATATGTATGCGCCCACTT 5437
Qy 1268 TGTGGGCAAGGATGATCTGATGACTCACTTCTTCTCCATCTTCTAGCCAGAGCAAC 1327
Db 5438 TGTGGGCAAGGATGATCTGATGACTCACTTCTTCTCCATCTTCTAGCTCAGGAACAAC 5497
Qy 1328 TTGAAGAGCCCTGATTTGTGAGTCTAGCGGGCTTGTACTCCATTTGAGCCACTTGACC 1387
Db 5498 TTGAAGAGCCCTAGATTTGTGAGATCTAGCGGGCGTGTACTCCATTTGAGCCACTTGACC 5557
Qy 1388 TACCTCAGATCATTTGAAGGACTCCAATGTCTTAGCGCAATTTTCACTCCATAGTTACTCTC 1447
Db 5558 TACCTCAGATCATTTCAAGCACTCCATGSCCTTAGGCAATTTTCACTCCATAGTTACTCTC 5617
Qy 1448 CAGGTGAGATCAATAGGTTGCTTTCATGCTCAGGAAGCTTTGGGGTACCACTTGGGAG 1507
Db 5618 CAGGTGAGATCAATAGGTTGCTTTCATGCTCAGGAAGCTTTGGGGTACCGGCTTGGGAG 5677
Qy 1508 TCTGAGACATCGGGCCAGAGTGTCCGGCTAGTACTGTCCAGGGGGGAGGGCGG 1567
Db 5678 TCTGAGACATCGGGCCAGAGTGTCCGGCTAGTACTGTCCAGGGGGGAGGGCGG 5737
Qy 1568 CCATTTGTGGCAAGTACTCTTCACTGGGAGTAAGCAACCACTTAACTCACTCCAA 1627
Db 5738 CCATTTGTGGCAAGTACTCTTCACTGGGAGTAAGCAACCACTTAACTCACTCCAA 5797
Qy 1628 TTCGGCTCGCTCCGGCTGGCACTTGTCCGGCTGCTTCTGCTGCTACAGCGGGGAG 1687
Db 5798 TCCGGCTCGCTCCAGTTCGATTTATCCAGTGTGCTGCTGCTTACAGCGGGGAG 5857
Qy 1688 ACATATATCACAGCCTGTCTGTGCCGACCCCGCTGGTTCATGTTGTCTACTCTCTAC 1747
Db 5858 ACATATATCACAGCCTGTCTGTGCCGACCCCGCTGGTTCATGTTGTCTACTCTCTAC 5917
Qy 1748 TCTCGTAGGGGTAGGCATCTATCTACTCTCCCAACCG 1784
Db 5918 TTCTGTAGGGGTAGGCATCTATCTACTCTCCCAACCG 5954

RESULT 9
US-10-434-842-16
; Sequence 16, Application US/10434842
; Publication No. US20040005549A1
; GENERAL INFORMATION:
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY RE
; FILE REFERENCE: 0342/IH3950S3
; CURRENT APPLICATION NUMBER: US/10/434,842
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 10/233,307
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US 10/005,469
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 16
; LENGTH: 7989
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: FCA4 Replicon Sequence
US-10-434-842-16

Query Match 87.9%; Score 1585; DB 17; Length 7989;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 8 GCATGTCAATGTCTTATACATGACAGGGCGCCTGATCACAACCGTGGCTGCGTGGAGAAA 67
Db 5975 GCTGCTCGATGTCTTACATATGACAGGGCGCCTGATCAGCGCATGCGTGGCGAGAAA 6034
Qy 68 GCAAGCTTGCCCATCAACGCGCTGAGCAACTCCCTTGTGCTGCGTCAACATAACCTTGGTCTTAT 127
Db 6035 CCAAGCTTGCCCATCAATGCACTGAGCAACTCTTGTCTCGTCAACACAACCTTGGTCTTATG 6094
Qy 128 CCAACAATCCCCAGTGCAGGCTTCGGCAGAGAAAGGTCACTTTTGAAGACTGCAAG 187
Db 6095 CTACACATCTCCGAGGCAAGCCTCGCGCAGAGAAAGGTCACTTTTGAAGACTGCAAG 6154
Qy 188 TCCTGACGATCAATTACCGGGAGCTGCTCAAGGAGATGAAGGGGAAGGGTCCACAGTGA 247
Db 6155 TCCTGACGACCACTACCGGGAGCTGCTCAAGGAGATGAAGGGGAAGGGTCCACAGTTA 6214
Qy 248 AGGCTAACTGCTATCTGTAGAGAGCATGCAAGCTGACGCCGCCGCCGCTTGGGCCAAAT 307
Db 6215 AGGCTAACTTCTATCCGTGGAGGAAGCCTGTAAAGCTGACGCCGCCGCCCATCTTCGCCAGAT 6274
Qy 308 CCAATTTGGCTATTTGGGCAAGGAGCTCGGAGCCTATCCAGCAGGGCGCTTAAACCA 367
Db 6275 CTAAATTTGCTATTTGGGCAAGGAGCTCGGAGCCTATCCAGCAAGGGCGTTAAACCA 6334
Qy 368 TCGCCTCCGTGTGGAGGACTTCTCTGGAGGACACTGACACACCAATTCAGACCAACATCA 427
Db 6335 TCCGCTCCGTGTGGAGGACTTCTCTGGAGGACACTGAGACACCAATTCAGACCAACATCA 6394
Qy 428 TGGCAAAAATAGGTTCTTCTGCTCAACACAGAGAAAGAGCGCCGCAACCACTGCGCC 487
Db 6395 TGGCAAAAATAGGTTCTTCTGCTCAACACAGAGAAAGAGGGGGCGCCGCAAGCCAGCTGCC 6454
Qy 488 TCATGCTATTTCCAGACCTGGGAGTTCTGTGTATGCGAGAGATGGCCCTTTACGAGCTGG 547
Db 6455 TTATGCTATTTCCAGATTTTGGGGTTCGTGTGTGCGAGAAAAATGGCCCTTTACGATGTGG 6514
Qy 548 TTTCCACTCTTCTCTAGGCGCTGATGGGCTCTCTCATACCGGATTTCCAATCTCTCTAAGC 607
Db 6515 TCTCCACCTCTCTCTAGGCGCTGATGGGCTCTTTCATACGGAATTTCCAATCTCTCTTGAC 6574
Qy 608 AGCGGTTCGAGTTCTTGTGTAATACCTTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 667
Db 6575 AGCGGTTCGAGTTCTTGTGTAATACCTTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 6634
Qy 668 ATGACACCGCTGTTTGTGACTCAACCGTCACTGAGATGACATCCGTTGTGAGGAGTCAA 727
Db 6635 ATGACACCGCTGTTTGTGACTCAACCGTCACTGAGAAATGACNTCCGTTGTGAGGAGTCAA 6694
Qy 728 TTTTACCAATGTGTGACTTGGCCCGCCGAGCTAGACAGGCGCATAAGGTCTGCTCACAGAGC 787
Db 6695 TCTACCAATGTGTGACTTGGCCCGCCGAGCAGACAGGCGCATAAGGTCTGCTCACAGAGC 6754
Qy 788 GGCTCTATGTGGGGGTCCCATGACTAACTCCAAAGGGCGAGAACTGCGGCTATCGCGGT 847
Db 6755 GGCTTTACATCGGGGGCCCCCTGACTAAATTTCTAAAGGGCAGAACTGCGGCTATCGCGGT 6814
Qy 848 GCGCGGAGCGGCTGCTGAGCTAGCTGCGGTAAATACCTTCACATGCTACTTGAAGG 907
Db 6815 GCGCGGAGCGGCTGACTGACGACCAAGCTGCGGTAAATACCTTCACATGTTACTTGAAGG 6874
Qy 908 CCGCTGCGACCTGTCTGAGCTGCCAAGCTCCAGGACTGACAGTCTCGTGAATGAGACG 967

```
Db 6875 CCGCTGGGCTCTCGAGCTGCGAAGCTCCAGGACTGCAGATGCTCGTATCGGAGAGC 6934
Qy 968 ACCTTGTCTTATCTGTGAAGCGCGGMAACCCAGAGAGCGCGGCAAGCCTACGAGTCT 1027
Db 6935 ACCTTGTCTTATCTGTGAAGCGCGGMAACCCAGAGAGCGCGGCAAGCCTACGAGTCT 6994
Qy 1028 TCACGGAGGCTATGACTAGTACTCTGCGCCCTCGGGGACCGCCCAACCGGATACG 1087
Db 6995 TCACGGAGGCTATGACTAGTACTCTGCGCCCTCGGGGACCGCCCAACCGGATACG 7054
Qy 1088 ACTTGGAGCTGATAACATCGTGTCTCCAAATGTGCGGTGCGCACAGATGATCTGGCA 1147
Db 7055 ACTTGGAGTGTATAACATCATGCTCTCCAAATGTGCGGTGCGCACAGATGATCTGGCA 7114
Qy 1148 AAGGGTGTATACCTCAACCGTACCCCAACCGTCCCTTCCGCGGGTGTGCGGAGA 1207
Db 7115 AAGGGTGTATATCTCAACCGTACCCCAACCGTCCCTTCCGCGGGTGTGCGGAGA 7174
Qy 1208 CAGCTAGGCACATCCAGTCAACTCTGCTAGGCAACATCATGATGATGCGCCACTT 1267
Db 7175 CAGCTAGGCACATCCAGTCAACTCTGCTAGGCAACATCATGATGATGCGCCACTT 7234
Qy 1268 TGTGGCAAGGATGATCTGATGACTCACTCTCTCTCCATCTCTTACCTAGCCACTTGACC 1327
Db 7235 TGTGGCAAGGATGATCTGATGACTCACTCTCTCTCCATCTCTTACCTAGCCACTTGACC 7294
Qy 1328 TTGAAAAGCCCTGGATTTGATGATCTAGGCTAGGCGGCTGTGTTACTTCCATTTAGCCACTTGACC 1387
Db 7295 TTGAAAAGCCCTAGATTTGATGATCTAGGCTAGGCGGCTGTGTTACTTCCATTTAGCCACTTGACC 7354
Qy 1388 TACCTCAGATCAATTGAACACTCCATGCTGTGTTAGGCGCAATTTTCACTCCATAGTTACTCTC 1447
Db 7355 TACCTCAGATCAATTGAACACTCCATGCTGTGTTAGGCGCAATTTTCACTCCATAGTTACTCTC 7414
Qy 1448 CAGGTGAGATCAATAGGCTGGCTTCAATGCTCAGGAGCTTGGGGTACACCCCTTGGAG 1507
Db 7415 CAGGTGAGATCAATAGGCTGGCTTCAATGCTCAGGAGCTTGGGGTACACCCCTTGGAG 7474
Qy 1508 TCTGGAGACATCGGGCCAGAGTTCGCGCTAAGTTACTGTCGCGGGGGGAGGCGG 1567
Db 7475 TCTGGAGACATCGGGCCAGAGTTCGCGCTAAGTTACTGTCGCGGGGGGAGGCGG 7534
Qy 1568 CCAATTGTGGCAAGTACTCTTCAACTGGGCGAGTAAAGCAACCAAGCTTAAACTCACTCCAA 1627
Db 7535 CCAATTGTGGCAAGTACTCTTCAACTGGGCGAGTAAAGCAACCAAGCTTAAACTCACTCCAA 7594
Qy 1628 TTCGGCTCGCTCCGCGCTGGACTTGTTCGGCTGTGTTGCTGCTACAGCGGGGAG 1687
Db 7595 TTCGGCTCGCTCCGCGCTGGACTTGTTCGGCTGTGTTGCTGCTACAGCGGGGAG 7654
Qy 1688 ACATATATACAGCTGTCTGTCGCGGACCCGCTGCTTCAATGTTGCTTACTCTCTAC 1747
Db 7655 ACATATATACAGCTGTCTGTCGCGGACCCGCTGCTTCAATGTTGCTTACTCTCTAC 7714
Qy 1748 TCTCGGTAGGGTAGGCATCTATCTACTTCCCAACCG 1784
Db 7715 TTTCTGTAGGGTAGGCATCTATCTACTTCCCAACCG 7751
```

RESULT 10

US-10-639-150-1

; Sequence 1, Application US/10639150

; Publication No. US20040121975A1

; GENERAL INFORMATION:

; APPLICANT: BRISTOL-MYERS SQUIBB COMPANY

; TITLE OF INVENTION: HEPATITIS B VIRUS ASSAYS

; FILE REFERENCE: D0224 NP

; CURRENT APPLICATION NUMBER: US/10/639,150

; CURRENT FILING DATE: 2003-08-12

; PRIOR APPLICATION NUMBER: US 60/402,661

; PRIOR FILING DATE: 2002-08-12

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 7989

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: HCV Replicon

US-10-639-150-1

Query Match 87.9%; Score 1585; DB 19; Length 7989;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

```
Qy 8 GCATGTCAATGTCTCTATACATGACAGCGCGCTGATCACCGTGCCTGCGGAGGAAA 67
Db 5978 GCTGCTCATGTCTTACACATGACAGCGCGCTCATCACGCTATGCGTGCAGGAAA 6037
Qy 68 GCNAGCTGCCATCAACCGGTGAGCAACTCTTCTGCTGCGTCAACCTGCTTATT 127
Db 6038 CCAAGCTGCCATCAATGCTGAGCAACTCTTCTGCTGCTGCTGCTGCTGCTGCTG 6097
Qy 128 CCACAAATCCCGAGTCAAGCTGCGGAGAGAGGTCAACCTTTTCAGAGACTGCAAG 187
Db 6098 CTACACATCTCGCAGCGCAAGCTGCGGAGAGAGGTCAACCTTTTCAGAGACTGCAAG 6157
Qy 188 TCCTGGAGCATCATTTACCGGAGCTGCTCAAGGAGATGAAGCGGAGGCTCCACAGTGA 247
Db 6158 TCCTGGAGCATCATTTACCGGAGCTGCTCAAGGAGATGAAGCGGAGGCTCCACAGTGA 6217
Qy 248 AGGCTAAACTGCTATCTGTAGAGAGCATGCAAGCTGACGCGCGCTGCTTGGGCAAT 307
Db 6218 AGGCTAAACTTCTATCCGTGGAGAGAGCTGTAAGCTGACGCGCGCTGCTTGGGCAAT 6277
Qy 308 CCAAAATTTGGCTATGGGCAAGAGCTGCGGAGCTATCCAGAGGCGGCTTAACCA 367
Db 6278 CTAATTTGGCTATGGGCAAGAGCTGCGGAGCTATCCAGAGGCGGCTTAACCA 6337
Qy 368 TCCGCTCGTGTGAAGGACTTGTGGAGGACACTGACACACCAATTCAGACACCACTCA 427
Db 6338 TCCGCTCGTGTGAAGGACTTGTGGAGGACACTGACACACCAATTCAGACACCACTCA 6397
Qy 428 TGGCAAAATAGAGTTTCTGCTGCCAACAGAGAGAGGCGGCAACAGCTCGCC 487
Db 6398 TGGCAAAATAGAGTTTCTGCTGCCAACAGAGAGAGGCGGCAACAGCTCGCC 6457
Qy 488 TCATCGTATCCAGAGCTTGGAGTGTGTATGCGAGAGATGCGCTTTCAGAGCTGG 547
Db 6458 TTATCGTATCCAGATTTTGGGGGTTGCTGTGTCGAGAAATGCGCTTTCAGAGTGG 6517
Qy 548 TTTCCACTCTTCCCTCAGCGGTGATGGCTCTCATACGGATTCGAATCTCTCTAAGC 607
Db 6518 TCTCCAGCTCCCTCAGCGGTGATGGCTCTCTATACGGATTCGAATCTCTCTAAGC 6577
Qy 608 AGCGGGTCCAGTCTCTGCTGAATACCTGGAAGCAAGAAATGCTTATGGCTTCTCAT 667
Db 6578 AGCGGGTCCAGTCTCTGCTGAATGCTTGAAGCAAGAAATGCTTATGGCTTCTCAT 6637
Qy 668 ATGACACCGCTGTTTGTGACTCAACGGTCACTGAGAAATGACATCGTGTGTGAGAGTCAA 727
Db 6638 ATGACACCGCTGTTTGTGACTCAACGGTCACTGAGAAATGACATCGTGTGTGAGAGTCAA 6697
Qy 728 TTTTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCGCATTAAGTTCGCTCAGAGC 787
Db 6698 TCTACCAATGTTGTGACTTGGCCCCCGAAGCTAGACAGGCGCATTAAGTTCGCTCAGAGC 6757
Qy 788 GGCCTTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGGGCTATCGCGGT 847
Db 6758 GGCCTTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTGGGCTATCGCGGT 6817
Qy 848 GCCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAATACCTCACAATGCTTGAAGG 907
Db 6818 GCCGCGGAGCGGCTGCTGACGACTAGCTGCGGTAATACCTCACAATGCTTGAAGG 6877
```



```
QY 908 CCCTGACGCTGTGAGTGTCCAAAGTCCAGACTGACAGATGCTCGTGAATGAGAGC 967
Db 6878 CCCTGCGGCTGTGAGTGTCCAAAGTCCAGACTGACAGATGCTCGTGAATGAGAGC 9937
QY 968 ACCTTGTCTTATCTGTGAAGCGCGGGAACCCAGAGAGCGCGGCAAGCTTACGAGTCT 1027
Db 6938 ACCTTGTCTTATCTGTGAAGCGCGGGAACCCAGAGAGCGCGGCAAGCTTACGAGTCT 6997
QY 1028 TCACGAGGCTATGACTAGTGTACTCTGCCCGCTGCGGACCCCGCCCAACCGGATACG 1087
Db 6998 TCACGAGGCTATGACTAGTGTACTCTGCCCGCTGCGGACCCCGCCCAACCGGATACG 7057
QY 1088 ACTTGAGCTGATPAACATCGTGTCTTCCAAATGTGCGTTCGACACAGATGATCTGGCA 1147
Db 7058 ACTTGAGTGTGATTAACATCATGCTCTCCAAATGTGCGTTCGACACAGATGATCTGGCA 7117
QY 1148 AAGGGTGTACTACTACCGGTGACCCACCGCTCCCGCTTGGCGGGCTGGTGGAGA 1207
Db 7118 AAGGGTGTACTACTACCGGTGACCCACCGCTCCCGCTTGGCGGGCTGGTGGAGA 7177
QY 1208 CAGTAGGCACACTCCAGTCACTCTGCTAGGTAAGGACATCATGTATGCGCCACTT 1267
Db 7178 CAGTAGGCACACTCCAGTCACTCTGCTAGGTAAGGACATCATGTATGCGCCACTT 7237
QY 1268 TGTGGCAAGGATGATCTCTGATGACTCACTTCTTCTCCATCTCTTAGCCAGGACAA 1327
Db 7238 TGTGGCAAGGATGATCTCTGATGACTCACTTCTTCTCCATCTCTTAGCCAGGACAA 7297
QY 1328 TTGAAAAGCCCTGGATGTTCAGATCTAGGGGCTTGTACTCCATTTAGCCACTTTGACC 1387
Db 7298 TTGAAAAGCCCTGGATGTTCAGATCTAGGGGCTTGTACTCCATTTAGCCACTTTGACC 7357
QY 1388 TACCTCAGATCAATGAAAGCTCCTAGTGTCTTAGGCGCATTTTCACTCCATAGTACTCTC 1447
Db 7358 TACCTCAGATCAATGAAAGCTCCTAGTGTCTTAGGCGCATTTTCACTCCATAGTACTCTC 7417
QY 1448 CAGGTAGATCAATAGGTTGGCTTCATGCTCAGGAGCTTGGGTACACCCCTTGGCGAG 1507
Db 7418 CAGGTAGATCAATAGGTTGGCTTCATGCTCAGGAGCTTGGGTACACCCCTTGGCGAG 7477
QY 1508 TCTGGAGACATCGGGCCAGAGTGTCCGGCTTAAGTTCATGCTCCAGCGGGGAGGGCG 1567
Db 7478 TCTGGAGACATCGGGCCAGAGTGTCCGGCTTAAGTTCATGCTCCAGCGGGGAGGGCG 7537
QY 1568 CCAATTTGTGGCAAGTACCTCTTCACTGGGCGAGTAGGACCAAGCTTAACTCACTCCAA 1627
Db 7538 CCAATTTGTGGCAAGTACCTCTTCACTGGGCGAGTAGGACCAAGCTTAACTCACTCCAA 7597
QY 1628 TTCCGGCTCGCTCCCGCTGAGCTTGTGCGGCTTGGTTCGTTGCTGCTACAGCGGGAG 1687
Db 7598 TTCCGGCTCGCTCCCGCTGAGCTTGTGCGGCTTGGTTCGTTGCTGCTACAGCGGGAG 7657
QY 1688 ACATATATCACAGCTGTCTGCTGCCGACCCGCTGGTTCATGTTGCTTACTCTTAC 1747
Db 7658 ACATATATCACAGCTGTCTGCTGCCGACCCGCTGGTTCATGTTGCTTACTCTTAC 7717
QY 1748 TCTCCGTAGGGTAGGCATCTATCTACTCCCAACCG 1784
Db 7718 TTTCTGTAGGGTAGGCATCTATCTACTCCCAACCG 7754
```

RESULT 12

```
US-10-005-469-1
; Sequence 1, Application US/10005469
; Publication No. US20020155133A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS Pharmaceuticals, Inc.
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REP
; FILE REFERENCE: 0342/1H3950S1
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245,866
```

```
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV replicon I377/NS3-3'UTR
US-10-005-469-1
```

Query Match 87.9%; Score 1585; DB 13; Length 7992;

Best Local Similarity 93.2%; Pred. No. 0;

Matches 1657; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

```
QY 8 GCATGTCMAATGTCTTATACATGACGAGCGGCGCTTATCACACCGTGCCTGCGGAGGAAA 67
Db 5978 GCTGTGATGTCTTACATGACATGACGAGCGGCGCTTATCACACCGTGCCTGCGGAGAAA 6037
QY 68 GCAGCTGCCATCAACCGGCTGACCAACTCTTCTGCTGCTCACCATTAACCTGCTTATT 127
Db 6038 CCAAGCTGCCATCAATGACCTGAGCAACTCTTCTGCTGCTCACCATTAACCTGCTTATG 6097
QY 128 CCACAAACATCCGCGAGTCAAGCCCTGCGGCAAGAGGTCACTTTTGACAGACTGCAAG 187
Db 6098 CTACAAACATCTCGACGCGCAAGCCCTGCGGCAAGAGGTCACTTTTGACAGACTGCAAG 6157
QY 188 TCCTGAGAGTCACTTACCGGCGAGCTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTGA 247
Db 6158 TCCTGAGAGTCACTTACCGGCGAGCTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTGA 6217
QY 248 AGGTAAACTGCTATCTGTAGAAGCATGCAAGCTGACGCGCCCGCTTTCGGCCAAAT 307
Db 6218 AGGTAAACTTCTATCCGTGGAGGAAAGCTGTAGCTGACGCGCCCGCTTTCGGCCAAAT 6277
QY 308 CCAATTTTGGCTATGCGGCAAGAGCTGCGGAGCTATCCAGAGCGCGCTTAAACCA 367
Db 6278 CCAATTTTGGCTATGCGGCAAGAGCTGCGGAGCTATCCAGAGCGCGCTTAAACCA 6337
QY 368 TCCGCTCGCTGGAAGGACTTGTCTGAGGACACTGACACACCAATTCAGACCCACCATCA 427
Db 6338 TCCGCTCGCTGGAAGGACTTGTCTGAGGACACTGACACACCAATTCAGACCCACCATCA 6397
QY 428 TGGCAAAATAGAGTTTCTGCGTCCAAACAGAGAAAGGAGCGCGCAACAGCTCGCC 487
Db 6398 TGGCAAAATAGAGTTTCTGCGTCCAAACAGAGAAAGGAGCGCGCAACAGCTCGCC 6457
QY 488 TCATGCTATCCAGACCTTGGAGTTCGTGTATGCGAGAAAGATGCGCTTTACGAGTGG 547
Db 6458 TTATGCTATCCAGATTTGGGGGTTGCTGTGTCGAGAAATGCGCTTTACGATGTTG 6517
QY 548 TTTCCACTTTCCTCAGCGCTGATGGCTCTCTATAGCGATTCATCTCTCTTAAGC 607
Db 6518 TTTCCACTTTCCTCAGCGCTGATGGCTCTCTATAGCGATTCATCTCTCTTAAGC 6577
QY 608 AGCGGTGAGTTCCTGCTGAATACCTGGAAGCAAGAAATGCGCTTATGGCTTCTCAT 667
Db 6578 AGCGGTGAGTTCCTGCTGAATACCTGGAAGCAAGAAATGCGCTTATGGCTTCTCAT 6637
QY 668 ATGACACCGCTGTTTTGACTCAACCGTCACTGAGAATGACATCCGTTTGAGGAGTCAA 727
Db 6638 ATGACACCGCTGTTTTGACTCAACCGTCACTGAGAATGACATCCGTTTGAGGAGTCAA 6697
QY 728 TTTTACCAGTGTGACTTGGCCCCCGAAGCTAGACAGGCCCATTAAGTTCGCTTCAAGAGC 787
Db 6698 TTTTACCAGTGTGACTTGGCCCCCGAAGCTAGACAGGCCCATTAAGTTCGCTTCAAGAGC 6757
QY 788 GCGCTTATGCTCGGGTCCCATGACTAACTCAAGGGCAGAACTGCGGCTATCGCGGT 847
Db 6758 GCGCTTATGCTCGGGTCCCATGACTAACTCAAGGGCAGAACTGCGGCTATCGCGGT 6817
QY 848 GCCGCGGAGCGGCTGCTGACGACTAGTTCGGGTAAATACCTCACAATGCTTGAAGG 907
Db 848 GCCGCGGAGCGGCTGCTGACGACTAGTTCGGGTAAATACCTCACAATGCTTGAAGG 907
```



```
QY 848 GCCCGCGAGCGCGCTGACGACTAGTTCGGTAAATACCCCTACATGCTACTTGAAGG 907
Db |||||
QY 6818 GCCCGCGAGCGGTGACTGACGACGAGCTGCGTAAATACCCCTACATGTTACTTGAAGG 6877
Db |||||
QY 908 CCCTGCGAGCTTCGAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAATGGAGAGC 967
Db |||||
QY 6878 CCCTGCGGCTTCGAGCTGCGAAGCTCCAGGACTGCACGATGCTCGTATGCGGAGAGC 6937
QY 968 ACCTTGTGTTATCTGTGAAGCGCGGAAACCCAGAGGACGCGGAAGCTTACGAGTCT 1027
Db |||||
QY 6938 ACCTTGTGTTATCTGTGAAGCGCGGAGACCCAGAGGACGAGGAGCTTACGCGCT 6997
QY 1028 TCACGAGGCTATGACTAGTACTCTGCGCCCTCGGGACCGCCCAACCGGATACG 1087
Db |||||
QY 6998 TCACGAGGCTATGACTAGTACTCTGCGCCCTCGGGACCGCCCAACCGGATACG 7057
QY 1088 ACTTGGAGCTGATAACATCGTGTCTCTCAATGTGCGGTGCGACACGATGCTTGGCA 1147
Db |||||
QY 7058 ACTTGGAGTTGATAACATCATGCTCTCCATGTGTGTCAGTGGGACGATGCTTGGCA 7117
QY 1148 AAGGCTGACTACTCCACCGGTGACCCCAACCGTCCCGCTGCGGGCTGCTGGGAGA 1207
Db |||||
QY 7118 AAGGCTGACTACTCTCACCGGTGACCCCAACCGTCCCGCTGCGGGCTGCTGGGAGA 7177
QY 1208 CAGCTAGGCACACTCCAGTCAACTCTGCTAGGCTAGGCAACATCATGATGCGCCACTT 1267
Db |||||
QY 7178 CAGCTAGACACACTCCAGTCAATCTCTGCTAGGCAACATCATGATGCGCCACTT 7237
QY 1268 TGTGGCAAGAGTATCTGATGACTCACTTCTCTCCATCTCTTCTAGCCAGGACCAAC 1327
Db |||||
QY 7238 TGTGGCAAGAGTATCTGATGACTCACTTCTCTCCATCTCTTCTAGCTTCTAGGCAAC 7297
QY 1328 TTGAAAGCCCTGGATGTGATGCTACGAGCTTACGGGCTTGTACTCCATGAGCCACTTGAC 1387
Db |||||
QY 7298 TTGAAAGCCCTAGATGTGATGCTACGAGGCTGTTACTCCATGAGCCACTTGAC 7357
QY 1388 TACCTCAGATCATTTCAACGACTCCATGGCTTAGCGCATTTTCACTCCATAGTTACTCTC 1447
Db |||||
QY 7358 TACCTCAGATCATTTCAACGACTCCATGGCTTAGCGCATTTTCACTCCATAGTTACTCTC 7417
QY 1448 CAGGTAGATCAATAGCTGGCTTCTGCTCAGGAGCTTGGGTACCACTTTCGCGAG 1507
Db |||||
QY 7418 CAGGTAGATCAATAGCTGGCTTCTGCTCAGGAGCTTGGGTACCACTTTCGCGAG 7477
QY 1508 TCTGGAGACATCGGGCCAGAGTGTCCGGCTAAGTTACTGTCCAGGGGGGAGGGCG 1567
Db |||||
QY 7478 TCTGGAGACATCGGGCCAGAGTGTCCGGCTAAGTTACTGTCCAGGGGGGAGGGCTG 7537
QY 1568 CCAATTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAAGCAAGCTTAAACTCACTCCAA 1627
Db |||||
QY 7538 CCAATTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAAGCAAGCTTAAACTCACTCCAA 7597
QY 1628 TTCGGCTCGTCCCGCTGGACTTGTTCGGCTGTTCTGTGCTGCTACAGCGGGGAG 1687
Db |||||
QY 7598 TCCGGCTCGTCCCGCTGGACTTGTTCGGCTGTTCTGTGCTGCTTACAGCGGGGAG 7657
QY 1688 ACATATATCACGCTGCTGTCGGCCGACCGGCTGTTCTATGTTGCTTACTTCTTAC 1747
Db |||||
QY 7658 ACATATATCACGCTGCTGTCGGCCGACCGGCTGTTCTATGTTGCTTACTTCTTAC 7717
QY 1748 TCTCCGTAGGGTAGGCATCTATCTACTCCCAACCG 1784
Db |||||
QY 7718 TTTCTAGGGTAGGCATCTATCTACTCCCAACCG 7754
```

RESULT 14

```
US-10-005-469-4
; Sequence 4, Application US/10005469
; Publication No. US2002015133A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS Pharmaceuticals, Inc.
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REH
```

```
; FILE REFERENCE: 0342/IH395US1
; CURRENT APPLICATION NUMBER: US/10/005,469
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Replicon RNA from cell line HCVR9
US-10-005-469-4
```

```
Query Match 87.9%; Score 1585; DB 13; Length 7992;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 8 GCATGTCATATGTCTTATACATGACAGCGGCCCTTGATCACACCGTGCCTGCGGAGGAAA 67
Db 5978 GCTGCTCGATGTCTTACATGACATGGACAGCGGCCCTTGATCAGCCATGCGTGGGAGGAAA 6037
QY 68 GCAAGCTGCCATCAACCGGCTGACCAACTCTTCTGCTGGTCACCATTAACCTGGTCTATT 127
Db 6038 CCAAGCTGCCATCAATCACTGACCAACTCTTCTGCTGGTCACCAACTTGGTCTATTG 6097
QY 128 CCACAAACATCCGCGAGTCAAGCCCTGCGGAGAAAGCTCACTTTTGACAGACTGCAAG 187
Db 6098 CTACAAACATCTCGCAGCGCAAGCCCTGCGGAGAAAGCTCACTTTTGACAGACTGCAAG 6157
QY 188 TCCTGGACGATCATTTACCGGAGCGTGTCTCAAGGAGATGAAGCGGAGCGGTCCACAGTGA 247
Db 6158 TCCTGGACGACCACTACCGGAGCGTGTCTCAAGGAGATGAAGCGGAGCGGTCCACAGTGA 6217
QY 248 AGGCTAAACTGCTATCTCTAGAAAGCATGCAAGCTGACGCCCGCCGATTCGGCCAAAT 307
Db 6218 AGGCTAAACTTCTATCCGTGGAGGAAGCTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6277
QY 308 CCAATTTGGCTATGGGCAAGGACGTCCGGAACCTTATCCAGCAAGGCGGTTAAACCA 6337
Db 6278 CTAATTTGGCTATGGGCAAGGACGTCCGGAACCTTATCCAGCAAGGCGGTTAAACCA 6337
QY 368 TCCGCTCGGTGGAAGGACTTGTCTGAGGACACTGACACACCAATTCAGACCAACATCA 427
Db 6338 TCCGCTCGGTGGAAGGACTTGTCTGAGGACACTGAGACACCAATTCAGACCAACATCA 6397
QY 428 TGGCAAAAATCAGGTTTTCTCGCTCCAAACAGAGAAAGGAGCGCGCAACAGCTCGCC 487
Db 6398 TGGCAAAAATCAGGTTTTCTCGCTCCAAACAGAGAGGGGGCGCAAGCGCTCGCC 6457
QY 488 TCATCGTATCCGAGACCTCGGAGTTCGTGTATGCGAGAAAGATGSCCTTTACGACGTGG 547
Db 6458 TTATCGTATCCGAGATTTGGGGGTTCTGTGTGCGAGAAATGSCCTTTACGATGCG 6517
QY 548 TTTCCACTCTTCTCAGGCGGTGATGGGCTCTCTCATACGATTCCTCAATCTCTCTCTAAGC 507
Db 6518 TCTCACCCTCTCTCAGGCGGTGATGGGCTCTTTCATACGATTCCTCAATCTCTCTCTGAC 6577
QY 608 AGCGGGTCCAGTTCCTCGTGAATACCTGCAAGCAAGAAATGCGCTATGGCTTCTCTCAT 667
Db 6578 AGCGGGTCCAGTTCCTCGTGAATGCTGGAAGCAAGAAATGCTCTATGGCTTCTCTCAT 6637
QY 668 ATGACACCCGCTGTTTTCGACTCAACCGGTCACTGAGAATGACATCCGTTTGAGGAGTCAA 727
Db 6638 ATGACACCCGCTGTTTTCGACTCAACCGGTCACTGAGAATGACATCCGTTTGAGGAGTCAA 6697
QY 728 TTTACCAATGTTGACTTTCGCTCCCGGAGCTAGACAGGCCATAGTTCGCTCAGAGC 787
Db 6698 TCTACCAATGTTGACTTTCGCTCCCGGAGCTAGACAGGCCATAGTTCGCTCAGAGC 6757
QY 788 GGCTCTATGTCGGGGTCCCATGACTACTTCAAGGGCAGACTCGCGGCTATCCCGGT 847
Db |||||
```

Db 6758 GGCATTACATCGGGGCCCTGACTAATTCTAAAGGCGACAACTGCGGCTATCGCGGT 6817
QY 848 CCGCGCGAGCGGCTGCTGAGCACTAGCTCGCGTAATACCTCACATGCTACTTGAAG 907
Db 6818 CCGCGCGAGCGGCTGCTGAGCACTAGCTCGCGTAATACCTCACATGCTACTTGAAG 6877
QY 908 CCGCTGCAAGCCTGCTGAGCTGCCAAGCTCCAGAGCTGCAGAGTCTCAGCATCTCTGTAATGGAGCG 967
Db 6878 CCGCTGCGGCTGCTGAGCTCGAAGCTCCAGAGCTGCACGATCTCGATGCGGAGCG 6937
QY 968 ACCTTGCTGTTATCTGTGAAGCGCGGAAACCCAGAGAGCGCGGCAAGCTACGAGTCT 1027
Db 6938 ACCTTGCTGTTATCTGTGAAGCGCGGAAACCCAGAGAGCGCGGCAAGCTACGAGCT 6997
QY 1028 TCAGGAGGCTATGACTAGGTACTCTGCCCCCTGCGGAGCCGCCCAACCGGAATACG 1087
Db 6998 TCAGGAGGCTATGACTAGATACTCTGCCGCCCTGCGGAGCCGCCCAACCGAATACG 7057
QY 1148 AAAGGCTGACTACTACCTACCGTGACCCACCGTCCCTTTGCGCGGCTGCGTGGGAGA 1207
Db 7118 AAAGGCTGACTATCTCACCGTGACCCACCGTCCCTTTGCGCGGCTGCGTGGGAGA 7177
QY 1208 CAGCTAGGCACACTCCAGTCAACTCTGCGTAGGCAACATCATCATGTATGCGGCCACTT 1267
Db 7178 CAGCTAGGCACACTCCAGTCAATTCTGCGTAGGCAACATCATGTATGCGGCCACTT 7237
QY 1268 TGTGGGCAAGGATGATCTGATGACTCATCTTCTCTCATCTTCTAGCCGAGGAGCAAC 1327
Db 7238 TGTGGGCAAGGATGATCTGATGACTCATCTTCTCTCATCTTCTAGCTCAGGAAACAA 7297
QY 1328 TTGAAGAGCCCTGGATTGTCAGATCTACGGGCTGTTACTCCATTGAGCCACTTGACC 1387
Db 7298 TTGAAGAGCCCTGATTGTCAGATCTACGGGCTGTTACTCCATTGAGCCACTTGACC 7357
QY 1388 TACCTCAGATCAATGAAGACTCCATGCTTTAGCGGATTTTCACTCCATGATTACTCTC 1447
Db 7358 TACCTCAGATCAATCAAGACTCCATGSCCTTAGGCAATTTCACTCCATGATTACTCTC 7417
QY 1448 CAGGTGAGATCAATAGGCTGCTTCATGCTCAGGAAGCTTGGGTTACCACTTGCAG 1507
Db 7418 CAGGTGAGATCAATAGGCTGCTTCATGCTCAGGAAGCTTGGGTTACCGCCCTTGCAG 7477
QY 1508 TCTGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTTACTGTCCAGGGGGAGGGCG 1567
Db 7478 TCTGAGACATCGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGAGGGCTG 7537
QY 1568 CCATTGTGGCAAGTACCTTCAACTGGGCGATGAAGGACCAAGCTTAACTCACTCCAA 1627
Db 7538 CCATTGTGGCAAGTACCTTCAACTGGGCGATGAAGGACCAAGCTTAACTCACTCCAA 7597
QY 1628 TTCGGCTGCGTCCCGCTGGACTTGTCCGCTGGTTTGTGCTGGCTACAGCGGGGAG 1687
Db 7598 TCCCGCTGCGTCCCGCTGGATTTATCCAGCTGGTTGCTGTTACAGCGGGGAG 7657
QY 1688 ACATATACAGGCTGTCTGTCGCCGACCCCGTGGTTCAATGTTGTCCTACTCTCTAC 1747
Db 7658 ACATATACAGGCTGTCTGTCGCCGACCCCGTGGTTCAATGTTGTCCTACTCTCTAC 7717
QY 1748 TCTCGGTAGGGTAGGCATCTACTCTCCGCAACCG 1784
Db 7718 TTCTGTAGGGTAGGCATCTACTCTCTCCGCAACCG 7754

RESULT 15

US-10-005-469-5
; Sequence 5, Application US/10005469
; Publication No. US2002015513A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS Pharmaceuticals, Inc.

; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFIC
; FILE REFERENCE: 0342/1H395US1
; CURRENT APPLICATION NUMBER: US/10/005,469
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Replicon from cell line HCVr22
US-10-005-469-5

Query Match 87.9%; Score 1585; DB 13; Length 7992;
Best Local Similarity 93.2%; Pred.No. 0;
Matches 1657; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 8 GCATGTCAATGCTCTATACATGACAGGCGCCCTGATCACACCGTGCCTGCGGAGGAAA 67
Db 5978 GCTGCTCGATGCTCTACACATGACAGAGCGCCCTGATCAGCCATGCGCTGCGGAGAAA 6037
QY 68 GCMAGCTGCCATCAACACGCGCTGAGCAACTCTCTGCTGCTGTCACCATAACTGCTCTATT 127
Db 6038 CCAAGCTGCCATCAATGCACTGAGCAACTCTTTGCTCGCTCACCACAACTTGGTCTATG 6097
QY 128 CCACACATCCCGCAGTGCAGCCCTGCGGAGAGAGTCACTTTCAGACACTGCGCAG 187
Db 6098 CTACACATCTCGACGCGAAGCTGCGGACAGAAAGTCACTTTTGACAGACTCGAG 6157
QY 188 TCGTCGACGATCAATTACCAGGAGCTGCTCAAGAGATGAAGCGAAGGCGTCCACAGTGA 247
Db 6158 TCTGAGACGACCACTACCGGAGCTGCTCAAGAGATGAAGGAGGCGTCCACAGTTA 6217
QY 248 AGGCTAACTGCTATCTGTAGAAGAGCATGCAAGCTGACGCGCCCGCAATTCGGCCAAAT 307
Db 6218 AGGCTAACTTCTATCCGTGGAGAGCTGTGAAGCTGACGCGCCCACTTCGGCCAGAT 6277
QY 308 CCAATTTGCTATGGGCAAGGAGCTCGGAGCTGTTCAGAGCTATTCAGCAGGCGGTTAACCA 367
Db 6278 CTAAATTTGCTATGGGCAAGGAGCTCGGAACTTATCCAGCAAGGCGGTTAACCA 6337
QY 368 TCCGCTCCGCTGGAAGGACTTCTGAGAGGACTGCAGACACTGCACACCAATTCAGACCACCATCA 427
Db 6338 TCCGCTCCGCTGGAAGGACTTCTGGAAGAGACTGCTGGAAGAGACTGCAGACCAATTTGACCCACCATCA 6397
QY 428 TGGCAAAAATGAGGTTTCTGCTCCAAACAGAGAAAGAGGCGCCGCAACACAGCTCGCC 487
Db 6398 TGGCAAAAATGAGGTTTCTGCTCCAAACAGAGAAAGGCGGCGCCGCAAGCTCGCC 6457
QY 488 TCATCGTATTCACAGACTTGGGAGTTGCTGTATGCGAAGATGCGCCCTTTACGAGCTGG 547
Db 6458 TTATCGTATTCACAGATTTGGGGGTTTCTGCTGCGAGAAAATGGGCCCTTACGATGTGG 6517
QY 548 TTTTCCACTCTTCTCAGGCGCTGATGGGCTCTCATACCGATTCCAATACTCTCTAAGC 607
Db 6518 TCTTCAACCTCTCCTCAGGCGGATGGGCTCTTATACCGATTTCCAATACTCTCTCTGGAC 6577
QY 608 AGCGGTCGAGTTCTCTGCTGAATACCTGGAAAGCAAGAAATGCCCCCTATGGGCTTCTCAT 667
Db 6578 AGCGGTCGAGTTCTTGGTGAATGCCCTGGAAGCGGAAGAAATGCCCCCTATGGGCTTCTCAT 6637
QY 668 ATGACACCCGCTGTTTTTGAATCAACCGTCACTGAGAAATGACATTCCTGTTTGAAGAGTCAA 727
Db 6638 ATGACACCCGCTGTTTTTGAATCAACCGTCACTGAGAAATGACATTCCTGTTTGAAGAGTCAA 6697
QY 728 TTTTCAAAATGTTGATCTTGGCCCCCGCAAGCTTAGACAGGCGCATAAAGGTCGCTCACAAGGC 787
Db 6698 TCTTACCAATGTTGATCTTGGCCCCCGCAAGCGCCACAGAGGCGCATAAAGGTCGCTCACAAGGC 6757

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2005, 13:25:15 ; Search time 308.897 Seconds
(without alignments)
9550.788 Million cell updates/sec

Title: US-10-712-479-1

Perfect score: 1803
Sequence: 1 acggtagcatgcaatgctc.....ggcatcaccatcaccatcac 1803

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1607.4	89.2	9472	4	Sequence 96, Appl
2	1596.2	88.5	9595	3	Sequence 4, Appli
3	1593	88.4	2991	1	US-08-014-416-4
4	1593	88.4	2991	1	US-08-324-977-49
5	1593	88.4	2991	2	US-08-384-616-49
6	1593	88.4	2991	2	US-08-904-686A-49
7	1593	88.4	2991	3	US-09-315-850-49
8	1593	88.4	7863	1	US-08-324-977-35
9	1593	88.4	7863	2	US-08-384-616-35
10	1593	88.4	7863	2	US-08-904-686A-35
11	1593	88.4	7917	1	US-09-315-850-35
12	1593	88.4	7917	1	US-08-324-977-31
13	1593	88.4	7917	2	US-08-384-616-31
14	1593	88.4	7917	2	US-08-904-686A-31
15	1593	88.4	9030	1	US-09-315-850-31
16	1593	88.4	9030	2	US-08-324-977-13
17	1593	88.4	9030	2	US-08-384-616-13
18	1593	88.4	9030	3	US-08-904-686A-13
19	1593	88.4	9416	1	US-09-315-850-13
20	1593	88.4	9416	2	US-08-324-977-1
21	1593	88.4	9416	2	US-08-384-616-1
22	1593	88.4	9416	3	US-08-904-686A-1
23	1593	88.4	9416	3	US-09-315-850-1
24	1585	87.9	7989	4	US-08-823-895A-27
25	1585	87.9	8001	4	US-09-539-601-10
26	1585	87.9	8001	4	US-09-539-601-7
27	1585	87.9	8637	4	US-09-539-601-22
					Sequence 4, Appli

28	1585	87.9	8649	4	US-09-539-601-13
29	1585	87.9	11076	4	US-09-539-601-1
30	1585	87.9	11076	4	US-09-539-601-25
31	1583.4	87.8	8001	4	US-09-539-601-28
32	1583.4	87.8	8638	4	US-10-029-907-6
33	1583.4	87.8	8638	4	US-10-029-907-7
34	1583.4	87.8	8638	4	US-10-029-907-24
35	1583.4	87.8	8638	4	US-10-029-907-25
36	1583.4	87.8	8639	4	US-10-029-907-1
37	1583.4	87.8	8642	4	US-10-029-907-2
38	1583.4	87.8	11076	4	US-09-539-601-31
39	1581.8	87.7	8001	4	US-09-539-601-16
40	1581.8	87.7	8643	4	US-10-029-907-4
41	1581.8	87.7	11076	4	US-09-539-601-19
42	1580.2	87.6	8648	4	US-10-029-907-5
43	1579.6	87.6	9413	4	US-09-827-688-6
44	1578.4	87.5	1773	4	US-09-720-035A-1
45	1574.8	87.3	1782	3	US-09-597-877-11

ALIGNMENTS

RESULT 1
US-08-150-204E-96
; Sequence 96, Application US/08150204E
; Patent No. 6538126
; GENERAL INFORMATION:
; APPLICANT: CHO, Joong Myung
; LEE, Yong Beom
; PARK, Young Woo
; LIM, Kook Jin
; CHOI, Deog Young
; SO, Hong Seob
; KIM, Chun Hyung
; KIM, Sung Taek
; YANG, Jae Young
; TITLE OF INVENTION: HEPATITIS C DIAGNOSTICS AND VACCINES
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YANG, Jae Young
; STREET: 386-1, Doryong-dong, Yuseong-gu
; CITY: Daejeon
; STATE: Daejeon
; COUNTRY: Republic of Korea
; ZIP: 305-340
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5inch, 1.44MB storage
; COMPUTER: IBM PC/pentium
; OPERATING SYSTEM: Windows
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/150,204E
; FILING DATE: 20-Apr-1994
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: KR 91-9510
; FILING DATE: 10-JUN-1991
; APPLICATION NUMBER: KR 91-13601
; FILING DATE: 6-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Shahan Islam, Esq.
; REGISTRATION NUMBER: 32,507
; REFERENCE/DOCKET NUMBER: 2695/FLK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 940-8564
; TELEFAX: (212) 940-8776
; INFORMATION FOR SEQ ID NO: 96
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9472 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: KHCV-LBCL1, Fig. 2
SEQUENCE DESCRIPTION: SEQ ID NO: 96
US-08-150-204E-96

Query Match 89.2%; Score 1607.4; DB 4; Length 9472;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY	8	GCATGTCATGTCCTATACATGGACAGGCGCCTGATCACACCCTGGCTGGCGAGGAAA	67
DB	7595	GCTGTTTCATGTCCTACACATGGACAGGCGCCTGATCAGCCATGGCTGGCGAGGAAA	7654
QY	58	GCAAGCTGCCCATCAACGCGCTGAGCACTCTCTGCTGGCTCACCATAACTGGTCTATT	127
DB	7655	GCAAGTTGCCCATCAACCGCTGAGCAATCTTTGCTACGTCACCAACATGCTCTATG	7714
QY	128	CCACAACATCCCGAGTGCAGAGCCTGCGGAGAGAAAGTCACTCTTTGACAGACTGCAAG	187
DB	7715	CTACAACATCCCGAGCGGCGCTGCGGAGAGAAAGTCACTCTTTGACAGACTGCAAG	7774
QY	188	TCCTGGAGCATCATTACCGGACCTGCTCAAGGAGATGAAGGCGAAGCGTCCACAGTGA	247
DB	7775	TCCTGGAGCACCTACCGGACCTGCTTAAGGAGATGAAGGCGAAGCGTCCACAGTGA	7834
QY	248	AGGCTAAACTGCTATCTGTAAGAGCATGCAAGCTGACGCCCCCGCATTCGGCCAAAT	307
DB	7835	AGGCTAAACTTCTATCTGTAGAAGAAAGCTGCAAACTGACGCCCCCACATTCGGCCAAAT	7894
QY	308	CCAAATTTGGCTATGGGCGAAGGACGTCGGAGCCTATCCAGCAGGCGCTTAACCCACA	367
DB	7895	CCAAATTTGGCTATGGGCGAAGGACGTCGGAGCCTATCCAGCAGGCGCTTAACCCACA	7954
QY	368	TCCGCTCCGCTGGAAGGACTTGCTGGAGGACACTGACACACCAATTCAGACCACTCA	427
DB	7955	TCGCTCTCGTGGAGGACCTGCTGGAGGACACTGAAACACCAATTAGCACATCACTCA	8014
QY	428	TGGCAAAAATAGAGTTTCTGCGTCCAAACAGAGAAAGGAGCGCGCAACACAGCTCGCC	487
DB	8015	TGGCAAAAATAGAGTTTCTGCTGTCCAAACAGAGAAAGGAGCGCGCAACACAGCTCGCC	8074
QY	488	TCATGCTATCCAGACCTGGAGTTCTGTATGCGAGAAATGGCCCTTTACAGCTGG	547
DB	8075	TTATCGCTTCCAGATCTGGAGTTCTGTATGCGAGAAATGGCCCTTTATGACGCTGG	8134
QY	548	TTTCCACATCTTCTCAGGCGGTGATGGCTCTCTCATACGGATTCCAATACTCTCTAAGC	607
DB	8135	TCTCCACCTTCTCAGGCGGTGATGGCTCTCTCATACGGATTCCAATACTCTCTAAGC	8194
QY	608	AGCGGCTCGAGTTCTGCTGTAATACCTGGAAGCAAGAAATGCCCTATGGGCTTCTCAT	667
DB	8195	AGCGGCTCGAGTTCTGCTGTAATACCTGGAATCAAGAAATGCCCTATGGGCTTCTCAT	8254
QY	668	ATGACACCGCTGTTTTCGACTCAAGCGTCACTGAGAAATGACATCCGTTTGAAGAGTCAA	727
DB	8255	ATGACACCGCTGTTTTCGACTCAACCGTCACTGAGAAATGACATCCGTTTGAAGAGTCAA	8314
QY	728	TTTACCAATGTTGACTTGGCCCCGAGCTAGACAGGCCATTAAGTCTGCTCAGAGC	787
DB	8315	TTTACCAATGTTGACTTGGCCCCGAGCTAGACAGGCCATTAAGTCTGCTCAGAGC	8374
QY	788	GGCTCTATGTGGGGGTCCTCATGACTCACTCAAGGCGCAAGCTGCGGCTATCGCCGT	847
DB	8375	GGCTCTATGTGGGGGTCCTCATGACTCACTCAAGGCGCAAGCTGCGGCTATCGCCGT	8434
QY	848	GGCGGCGGAGCGGCTGCTGAGCTAGCTCGGTAATACCCCTCACAATGCTACTTGAAG	907
DB	8435	GGCGGCGGAGCGGCTGCTGAGCTAGCTCGGTAATACCCCTCACAATGCTACTTGAAG	8494

QY	968	ACCTTCTCCTTATCTGTGAAAGCGGGGAAACCAAGAGGACGCGGCAAGCCTACAGTCT	1027
DB	8555	ACCTTCTCCTTATCTGTGAAAGCGGGGAAACCAAGAGGATCGCGGAGCCTACAGTCT	8614
QY	1028	TCACGAGGCTATGACTAGTACTCTGCCCGCCCTGGGGACCGGCCCAACCGGAATACG	1087
DB	8615	TCACGAGGCTATGACTAGTACTCTGCCCGCCCTGGGGACCGGCCCAACCGGAATACG	8674
QY	1088	ACTTGGAGCTGATTAACATCGTGTCTCCAATGTGTCGGTCCGACACATGATCTGGCA	1147
DB	8675	ACTTGGAGCTGATTAACATCATGTCTCCAATGTGTCGGTCCGACACATGATCTGGTA	8734
QY	1148	AAAGGCTGTACTACCTCACCCGCTGACCCACCGTCCCCCTTGGCGGGCTGGGTGGAGA	1207
DB	8735	AAAGGCTGTACTACCTCACCGCTGACCTACCAACCCCTTGCACGGGCTGGGTGGAGA	8794
QY	1208	CAGCTAGGCACACTCCAGTCAACTCTGTGCTAGGCAACATCATGATGCGCCACCTT	1267
DB	8795	CAGCTAGACACACTCCAGTCAACTCTGTGCTAGGCAACATCATGATGCGCCACCT	8854
QY	1268	TGTGGCAAGGATGATTTCTGATGACTCACTTCTTCTCCATCTTCTAGCCAGGACAC	1327
DB	8855	TATGGCAAGGATGATTTCTGATGACTCACTTCTTCTCCATCTTCTAGCTAGGACAA	8914
QY	1328	TTGAAAAGCCCTGGATTGTCAGATCTACGGGCTTGTACTTCCATTGAGCCACTTGACC	1387
DB	8915	TTGAAAAGCCCTTAGATTGTCAGATCTACGGGCTTGTACTTCCATTGAAACCACTTGATC	8974
QY	1388	TACTCAGATCATTTGAACGACTCCATGCTTTAGGCGCTTTTCACTCCATAGTTACTCTC	1447
DB	8975	TACTCAGATCATTTGAGCGACTCCATGGTCTTAGCGCATTTTCACTCCATAGTTACTCTC	9034
QY	1448	CAGGTGAGTCAATAGGCTGGCTTCATGCTCAGGAGCTTGGGGTACCACTTTCGGAG	1507
DB	9035	CAGGCGAGATCAATAGGCTGGCTTCATGCTCAGAAAACTTGGGGTACCACTTTCGGAG	9094
QY	1508	TCTGGAGACATCGGGCCAGAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGCGCG	1567
DB	9095	CTTGGAGACATCGGGCCAGAGTGTCCGCGCTAAGTTACTGTCCAGGGGGGAGGCGCG	9154
QY	1568	CCATTTGTGGCAAGTACCTCTTCAACTGGGCGACTAAGGACCAAGCTTAAACTCCTCAA	1627
DB	9155	CCATTTGTGGCAAGTACCTCTTCAACTGGGCGGTGAGGACCAAGCTTAAACTCCTCAA	9214
QY	1628	TTCGCGCTGCGTCCCGCTGAGACTTGTCCGGCTGGTTCTGCTGCTGCTACGCGGGAG	1687
DB	9215	TCCAGCGCGCTCCCGGTTGGACTTGTCCGGCTGGTTCTGCTGCTGCTACGCGGGAG	9274
QY	1688	ACATATATCAGACCTGCTCTGTCGCGGACCCCGCTGGTTCATGTTGCTGCTACTCTAC	1747
DB	9275	ACATATATCAGACCTGCTCTGTCGCGGACCCCGCTGGTTCATGTTGCTGCTACTCTAC	9334
QY	1748	TCTCCGTAGGGGTAGGCTATCTACTCTCCCAACCG	1784
DB	9335	TTTCCGTGGGGTAGGCTATCTACTCTCCCAACCG	9371

RESULT 2
US-09-014-416-4
; Sequence 4, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Buhk, Jens
; APPLICANT: Emerson, Susanne U.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; TITLE OF INVENTION: USPS THEREOF

```

; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 9595
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-014-416-4

```

Query Match 88.5%; Score 1596.2; DB 3; Length 9595;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 113; Indels 0;

		Matches 1664;	Conservative	0;	Mismatches	113;	Indels	0;	Gaps	0
Qy	8	GCATGTCAATGTCCTATACATGGACAGCGGCGCTGATCACACCGTGCCTCGGAGGAAA	67							
Db	7594	GCTGCTCAATGTCCTATACGTGGACAGCGGCGCTGATCAGCGCATCGCGTCGGAGGAAA	7653							
Qy	68	GCAAGCTGCCCATCAACCGCGTGAGCAACTCTCTTGCTGCGGTACCATAACTTGTCTATT	127							
Db	7654	GTAAGCTGCCCATCAACCGGTGAGCAACTCTTGCTGCGGTACCATAACTTGTCTACG	7713							
Qy	128	CCACAACATCCCGCAGTGCACAGCCCTCGGCGCAGAGAAAGGTCCACTTTTGACAGACTGCAAG	187							
Db	7714	CCACAACATCCCGCAGCGCAAGCCCTCGGCGCAGAGAGGTCCACTTTTGACAGATTGCAAG	7773							
Qy	188	TCCTGGAGATCATATTACCGGACGTGTCTCAAGGAGATGAAGCGCAAGCGCTCCACAGTGA	247							
Db	7774	TCCTGGATGATCATATTACCGGACGTCTCAAGGAGATGAAGCGCAAGCGGTCCACAGTTA	7833							
Qy	248	AGGCTAAACTGCTATCTGTAGAAAGCATGCAAGCTGACGCCGCCCGGATTCGGGCGCAAA	307							
Db	7834	AGGCTAAAGCTTCATCTATAGAGAGCGCTGCAAGCTGACGCCGCCCACTTCGGGCGCAAA	7893							
Qy	308	CCAAATTTGGCTATGGGCGCAAGACGTCGGAGCGCTATCCAGCAGAGCGCGTTAACCA	367							
Db	7894	CCAAATTTGCTATGGGCGCAAGAGCGTCGGAACCTATCCAGCAGAGCGCGTTAACCA	7953							
Qy	368	TCGCTCCGTGTGAAGGACTTGTCTGAGGACACTGTACACACCAATTTACAGACCACTCA	427							
Db	7954	TCGCTCCGTGTGGGAGGACTTGTCTGAAGACACTGTAAACACCACTTTACAGACCACTCA	8013							
Qy	428	TGCGAAAAATGAGTTTTCGCGTCCAAACGAGAGAAAGGCGCGCAAAACCAAGCTCGCC	487							
Db	8014	TGCGAAAAAGTGAGTTTTCGCGTCCAAACGAGAGAAAGGCGCGCAAAACCAAGCTCGCC	8073							
Qy	488	TCATCGTATTCCCAAGACTCGGAGTTCGTGTATGCGAGAGAGTGGCCCTTTACGACGTGG	547							
Db	8074	TTATCGTATTCCCAAGACTCGGAGTTCGTGTATGCGAGAGAGTGGCCCTTTACGACGTGG	8133							
Qy	548	TTTCCACTCTTCCTCAGGCGGTGATGGGCTCCTCATACGGATTCCTCAATCTCTCCTAAGC	607							
Db	8134	TTCTCCACCTTCCTCAGGCGGTGATGGGCTCCTCATACGGATTCCTCAATCTCTCCTAAGC	8193							
Qy	608	AGCGGTCGAGTTCTTGTTGAATACCTTGAAGAGCAAGAAATGCGCTTATGGGCTTCTCAT	667							
Db	8194	AGCGGTCGAGTTCTTGTTGAATACCTTGAAGAAATGCGCTTATGGGCTTCTCAT	8253							
Qy	668	ATGACACCCGCTGTTTTGACTCAACGCTCACTGAGATGACATCCGTTGTTGAGAGTCAA	727							
Db	8254	ATGACACCCGCTGTTTTGACTCAACGCTCACTGAGATGACATCCGTTGTTGAGAGTCAA	8313							
Qy	728	TTTTACCAATGTTGTGACTTGGCCCCGGAAGCTAGACAGGCGCATAAAGTGCCTCACAGAGC	787							
Db	8314	TTTTACCAATGTTGTGACTTGGCCCCGGAAGCGCAGACAGGCGCATAAAGTGCCTCACAGAGC	8373							
Qy	788	GGCTCTATGTGCGGGGTCCCATGACTAACTCCAAAAGGCGCAAACTCGCGGTATCGCCGGT	847							
Db	8374	GGCTTTTACATCGGGGTCCCTCAGCTAACTCAAAAGGCGCAAGACTCGCGGTATCGCCGGT	8433							
Qy	848	GCGCGCGGAGCGCGTCTGACGACTAGCTGCGGTAAATACCTTCATGCTACTTGAAGG	907							
Db	8434	GCGCGCAAGTGCCTGCTGACGACTAGCTGCGGTAAATACCTTCATGCTACTTGAAGG	8493							

QY	908	CCGCTG	CAGCCTG	TCGAGCTGCC	AAAGCTCC	AAGGACTC	ACGATCCT	CGTGAATGG	AAGC	967	
Db	8494	CCACTG	CAGCCTG	TCGAGCTGCA	AAAGCTC	CAGGACTC	ACGATCCT	CGTGAACGG	AAGC	8553	
QY	968	ACCTTGT	CGTTATCTGT	GAAAGCGCGG	GAACCCCA	AGGAGCGCG	GAAGCCT	TACGAGCTCT	1027		
Db	8554	ACCTTGT	CGTTATCTGT	GAAAGCGCGG	AACCCAG	AGGAGTGC	CGCGCGCCT	TACGAGCCT	8613		
QY	1028	TCACGG	AGGCTATG	ACTAGT	AGTACTCT	CGCCCCCT	CTGGGGAC	CGCCCCAAC	CGGAATACG	1087	
Db	8614	TCACGG	AGGCTATG	ACTAGT	ATATCC	CGCCCCCT	CTGGGGAT	TCGCCCCAAC	CAGAAATACG	8673	
QY	1088	ACTTGG	AGCTGAT	AAACATCG	TGTTCT	CCCAATGT	GTGCGT	CGCACAGAT	GCATCTGGCA	1147	
Db	8674	ACTTGG	AGCTGAT	AAACATCA	TGTTCT	CCCAATGT	GTGCGT	CGCGACGAT	GCATCTGGCA	8733	
QY	1148	AAAGGG	TGTACT	CTCAC	CGTGAC	CCCAAC	CGTCC	CCCTTGG	CGCGGCTG	CGTGGAGA	1207
Db	8734	AAAGGG	TATACT	CTCAC	CGTGAC	CCCAAC	CGTCC	CCCTTGG	CGCGGCTG	CGTGGAGA	8793
QY	1208	CAGCT	AGGCACACT	CCAGTCA	ACTCT	CTGGCTAG	GCACAT	CATCAT	TGATG	CGCCCACTT	1267
Db	8794	CAGCT	AGGCACACT	CCCAATCA	ACTCT	TGGCTAG	GCACAT	CATCAT	TGATG	CGCCCACTT	8853
QY	1268	TGTGGG	CAAGGATG	ATTCTG	ATGACT	CACATTTCT	CCCATCT	TCCTTAG	CCACAGG	ACCAAC	1327
Db	8854	TATGGG	CAAGGATG	ATTCTG	ATGACT	CACATTTCT	TCCCATCT	TCCTTAG	CCATCAAG	ACCAAC	8913
QY	1328	TTGAAA	AAAGCCCTG	GATTTG	CAGATCT	TACGGG	GCCTTGT	TACTCCAT	TGAGC	CACTTGGACC	1387
Db	8914	TTGAAA	AAAGCCCTG	GATTTG	CAGATCT	TACGGG	GCCTTGT	TACTCCAT	TGAGC	CACTTGGACC	8973
QY	1388	TACCT	CAGATCAT	TTGAAC	GACTCC	ATGCTTTAG	CGGATTTT	CACTCC	ATAGTTACT	CTCTC	1447
Db	8974	TACCT	CAGATCAT	TTGAAC	GACTCC	ATGCTTTAG	CGGATTTT	CACTCC	ATAGTTACT	CTCTC	9033
QY	1448	CAGGT	GAGATCAAT	ATGGT	GGCTTC	ATGCTC	TAGGAAGCT	TGGGGT	ACCAC	CCCTTGGCAG	1507
Db	9034	CAGGT	GAGATCAAT	ATGGT	GGCTTC	ATGCTC	TAGGAAC	CTTGGGGT	ATACC	CCCTTGGCAG	9093
QY	1508	CTTGG	AACATCGGG	CCAGAAGT	TGTCGG	CGCTAA	GTTACTGT	CTCCAGG	CGGGAG	CGGCG	1567
Db	9094	CTTGG	AACATCGGG	CCAGAAGT	TGTCGG	CGCTAA	GTTACTGT	CTCCAGG	CGGGAG	CGGCG	9153
QY	1568	CCATTT	TGTGCAAGT	ACTCTT	CAACTGG	GCAGT	AAGSAC	CAAGCTTAA	AACTCACT	CCCAA	1627
Db	9154	CCATTT	TGTGCAAGT	ACTCTT	TAAC	TGGGCAGT	TAAGSAC	CAAGCTTAA	AACTCACT	CCCAA	9213
QY	1628	TTCCGG	CTCGCTCC	CGGCTG	GAATG	TCGCGT	GTGTTG	TCTGCT	GGCTAC	CACCGGGGAG	1687
Db	9214	TTCCGG	CTCGCTCC	CGGCTG	GAATG	TCGCGT	GTGTTG	TCTGCT	GGCTAC	CACCGGGGAG	9273
QY	1688	ACATAT	ATCAC	AGCCTGTCT	CGTGG	CCGACCC	CGCTGG	TTCAT	TTGTG	CTTACTCTCTAC	1747
Db	9274	ACATAT	ATCAC	AGCCTGTCT	CGTGG	CCGACCC	CGCTGG	TTCAT	TTGTG	CTTACTCTCTAC	9333
QY	1748	TCTCCG	TAGGGTAG	GCATCTAT	CTACT	CCCCAA	CCG	1784			
Db	9334	TTTCTG	TAGGGTAG	GCATTTT	ACTTGT	CTCTCT	CCCCAA	CCG	9370		

RESULT 3

RESULT 32
 US-08-324-977-49
 ; Sequence 49, Application US/08324977
 ; Patent No. 5,747,319
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao
 ; APPLICANT: MORI, Chiatsu
 ; APPLICANT: TAKAMIZAWA, Aka-hisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B
 ; TITLE OF INVENTION: CDNA AND ANT

2414 CAGCTAGACACACTCCAGTTAACTCTCTGGCTAGGCAACATTATATGTATGCGCCACTT 2473
QY 1268 TGTGGGCAAGGATGATCTGTAGTACACATCTCTTCCATCTCTTACGCGCAGGAGCAAC 1327
Db 2474 TGTGGGCAAGGATGATCTGTAGTACACATCTCTTCCATCTCTTACGCGCAGGAGCAAC 2533
QY 1328 TTGAAAAAGCCCTGGGATGTGAGATCTAGGGGCTGTGTACTCCATTTAGGACCACTTGACC 1387
Db 2534 TTGAAAAAGCCCTGGGATGTGAGATCTAGGGGCTGTGTACTCCATTTAGGACCACTTGACC 2593
QY 1388 TACCTCAGATCAATGAAAGCACTCCATGTCTTGTAGCGCATTTTCACTCCATAGTACTCTC 1447
Db 2594 TACCTCAGATCAATGAAAGCACTCCATGTCTTGTAGCGCATTTTCACTCCATAGTACTCTC 2653
QY 1448 CAGGTGAGATCAATAGGGTGGCTTTCATGTCTCAGGAAGCTTGGGGTACCACTTGGCGAG 1507
Db 2654 CAGGTGAGATCAATAGGGTGGCTTTCATGTCTCAGGAAGCTTGGGGTACCACTTGGCGAG 2713
QY 1508 TCTGGAGACATGGGCGCAGAGTGTCCGCGCTAAGTACTGTCCAGGGGGGAGGGCG 1567
Db 2714 TCTGGAGACATGGGCGCAGAGTGTCCGCGCTAAGTACTGTCCAGGGGGGAGGGCG 2773
QY 1568 CCAATTGTGCGCAAGTACCTCTTCACTGGGCGAGTAAGCAAGCTTAACTCACTCCAA 1627
Db 2774 CCAATTGTGCGCAAGTACCTCTTCACTGGGCGAGTAAGCAAGCTTAACTCACTCCAA 2833
QY 1628 TTCCGGCTCGTCCCGCTGGGATGTCTCGGCTGGTTCGTCTGGCTACAGCGGGGAG 1687
Db 2834 TTCCGGCTCGTCCCGCTGGGATGTCTCGGCTGGTTCGTCTGGCTACAGCGGGGAG 2893
QY 1688 ACATATATCAGACCTGTCTCGTGCCGACCCCGCTGGTTCATGTGTGCTTACTCTAC 1747
Db 2894 ACATATATCAGACCTGTCTCGTGCCGACCCCGCTGGTTCATGTGTGCTTACTCTAC 2953
QY 1748 TCTCGTAGGGTAGCATCTATCTACTTCCCAACCG 1784
Db 2954 TTTCTGTAGGGTAGCATCTACTCTCTCCCAACCG 2990

RESULT 4

US-08-384-616-49
; Sequence 49, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hitoto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
; ADDRESS: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466

; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna from genomic RNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2991
; OTHER INFORMATION: /note: "sequence = 6372 - 9362 of
; OTHER INFORMATION: SEQ ID NO: 1"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2991
; US-08-384-616-49
Query Match 88.4%; Score 1593; DB 2; Length 2991;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 1662; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 8 GCATGTCAATGCTCTATACATGGACGCGCCCTGATCACAACCGTGCCTCGGAGGAAA 67
Db 1214 GCTGCTCAATGCTCTATACATGGACGCGCCCTGATCACAACCGTGCCTCGGAGGAAA 1273
QY 68 GCAAGCTGCCCATCAACGCGCTGAGCACTCTTGTGCTGCGTCACCATACCTGCTCTATT 127
Db 1274 GCAAGCTGCCCATCAACGCGCTGAGCACTCTTGTGCTGCGTCACCATACCTGCTCTATT 1333
QY 128 CCACAACATCCCGCAGTGCAGGCTGCGGCGAGAGAGTCCACCTTTGACAGACTGCAAG 187
Db 1334 CCACAACATCCCGCAGGCGGCGCTGCGGCGAGAGAGTCCACCTTTGACAGACTGCAAG 1393
QY 188 TCCTGGAGATCATTTACCGGACGCTCTCAAGGAGATGAAGCGAAGCGTCCACAGTGA 247
Db 1394 TCCTGGAGATCATTTACCGGACGCTCTCAAGGAGATGAAGCGAAGCGTCCACAGTGA 1453
QY 248 AGGCTAAACTGCTATCTGTAGAAGAGATGCAAGCTGACGCCCCCGCATTCGGCCCAAT 307
Db 1454 AGGCTAAACTGCTATCTGTAGAGAGCTGACGCCCCCGCATTCGGCCCAAT 1513
QY 308 CCAATTTGGCTATGGGCGAAGGACCTCCGAGGCTTATCCAGAGGCGCGTTAAGCACA 367
Db 1514 CCAAGTTTGGCTATGGGCGAAGGACCTCCGAGGCTTATCCAGAGGCGCGTTAAGCACA 1573
QY 368 TCCGCTCCGTGTGGAAGGACTTGTCTGGAGGACACTGACACACCAATTTGACAGCACCATCA 427
Db 1574 TCCACTCCGTGTGGAAGGACTTGTCTGGAGGACACTGACACACCAATTTGACAGCACCATCA 1633
QY 428 TGGCAAAAAATGAGGTTTCTGCTCCAAACAGAGAGAGGCGCGCAAAACAGCTCGCC 487
Db 1634 TGGCAAAAAATGAGGTTTCTGCTCCAAACAGAGAGAGGCGCGTAAAGCAGCCCGCC 1693
QY 488 TCATCGTATTTCCAGACCTGGGAGTTCGTGTATCGGAGAGATGCGCCCTTTACACGCTGG 547

Db 1694 TTATCGTATCCAGATCTGGAGTCCGTGTATCGGAGAGATGGCCCTCTATGATGTGG 1753
Qy 548 TTTCCACTCTTCTCAGGCCGTGATGGCTCCTCATACGGATTCCAATACTCTCTTAAGC 607
Db 1754 TCTCCACCCTTCTCAGCTCGTATGGCTCCTCATACGGATTCCAGTACTCTCTCGGC 1813
Qy 608 AGCGGTCCAGTCTCTGTGAATACCTGGAAGCAAGAAATAGCCCTATGGCTTCTCAT 667
Db 1814 AGCAGTCCAGTCTCTGTGAATACCTGGAATCAAAAGAAACCCCATGGCTTTTCAT 1873
Qy 668 ATGACACCGCTGTTTTCAGTCAACGGTCACTGAGATGACATCCGTGTTGAGGAGTCAA 727
Db 1874 ATGACACTCGCTGTTTTCAGTCAACGGTCAACGAGACGACATCCGTGTTGAGGAGTCAA 1933
Qy 728 TTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCCCATTAAGGTGCTCACAGAGC 787
Db 1934 TTTACCAATGTTGACTTGGCCCCCGAAGCTAGACAGGCCCATTAAGGTGCTCACAGAGC 1993
Qy 788 GGCTCTATGTCGGGGTCCCATGACTAACTCCAAAGGCGAGAACTCGGGCTATCGCCGGT 847
Db 1994 GGCCTTATATCGGGTCTCTGACTTAATCAAAAGGCGAGAACTCGGGTTATCGCCGGT 2053
Qy 848 GCCGGCGAGCGGTGCTGACGACTAGCTCGGTAAATACCTCACATGCTACTTTGAAG 907
Db 2054 GCCGGCGAGCGGTGCTGACGACTAGCTCGGTAAACCCCTCACATGTTACTTTGAAG 2113
Qy 908 CGCTGACGCTGTCGAGCTGCCAAGCTCCAGGACTCGAGCATGCTCGTGAATGAGAGC 967
Db 2114 CCTTCGAGCTGTCGAGCTCGAAGCTCCAGGACTCGACATGCTCGTGAACGAGAGC 2173
Qy 968 ACCTGTCTTATCTGTGAAAGCGCGGAACCCAAAGAGAGCGCGCAAGCTACGAGTCT 1027
Db 2174 ACCTGTCTTATCTGTGAAAGCGCGGAACCCAAAGAGAGCGCGCGAGCTACGAGTCT 2233
Qy 1028 TCACGAGGCTATGACTAGTACTCTGCCGCCCTTGGGGACCCGCCCAACCGGAATACG 1087
Db 2234 TCACGAGGCTATGACTAGTACTCGGCCGCCCTTGGGGACCCGCCCAACCGGAATACG 2293
Qy 1088 ACTTGAGCTGATAACATCGTGTCTCTCATGTGTGGTGGCAACGATGATCTGGCA 1147
Db 2294 ACTTGAGCTGATAACATCATGTTCTCCAAATGTGTGGTGGCAACGATGATCTGGCA 2353
Qy 1148 AAAGGGTGTACTACTCTACCGTACCCCGTCCCTTGGCGGGCTCGGTGGGAGA 1207
Db 2354 AAAGGGTGTACTACTCTACCGTATCCCAACCCCTTAGCAGGGCTCGGTGGGAGA 2413
Qy 1208 CAGCTAGGACACTCCAGTCAACTCTGGCTAGGCAACATCATGATGATCGGCCACTT 1267
Db 2414 CAGCTAGACACTCCAGTCAACTCTGGCTAGGCAACATCATGATGATGATCGGCCACTT 2473
Qy 1268 TGTGGGCAAGATGATCTGATGACTCATCTTCTTCATCTCTTAGCCAGGAGCAAC 1327
Db 2474 TGTGGGCAAGATGATCTGATGACTCATCTTCTTCATCTCTTAGCCAGGAGCAAC 2533
Qy 1328 TTGAAAAACCCCTGATTTGTGATCTACGGGCTTGTACTCCATGAGCACTTGACC 1387
Db 2534 TTGAAAAACCCCTGACTGCCAGATCTACGGGGCTGTTACTCCATGAGCACTTGACC 2593
Qy 1388 TACCTCAGATCATGCAACGACTCCATGTCTTAGGCAATTTTCACTCCATGATTTACTCTC 1447
Db 2594 TACCTCAGATCATGCAACGACTCCATGTGCTTAGGCAATTTTCACTCCATGATTTACTCTC 2653
Qy 1448 CAGGTGAGATCAATAGGGTGGCTTTCATGCTCAGGAAGCTTGGGTACCACTTTGCGAG 1507
Db 2654 CAGGTGAGATCAATAGGGTGGCTTTCATGCTCAGGAAGCTTGGGTACCACTTTGCGAG 2713
Qy 1508 TGTGAGACATCGGGCAGAAAGTGTCCGGCTTAAGTTACTGTCCAGGGGGGAGGGCG 1567
Db 2714 TGTGAGACATCGGGCAGAGAGCTCCGGCTTAGGCTACTGTCCAGGGGGGAGGGCG 2773
Qy 1568 CCATTGTGGCAAGTACTCTTCAACTGGGAGTAAGGACCAAGCTTAACTCACTCCAA 1627
Db 2774 CCATTGTGGCAAGTACTCTTCAACTGGGAGTAAGGACCAAGCTTAACTCACTCCAA 2833

Qy 1628 TTCGGCTCGTCCCGCTGGACTGTCCGGCTGGTTCGTTGCTGGCTACAGGGGGAG 1687
Db 2834 TCCGGCTCGTCCCGCTGGACTGTCCGGCTGGTTCGTTGCTGGCTACAGGGGGAG 2893
Qy 1688 ACATATATCACAGCTGTCTCGTGCCGACCCCGCTGTTCAATGTTGCTGCTACTCTTAC 1747
Db 2894 ACATATATCACAGCTGTCTCGTGCCGACCCCGCTGTTGTTCAATGTTGCTGCTACTCTTAC 2953
Qy 1748 TCTCCGTAGGGTAGGCATCTATCTACTCCCAACCG 1784
Db 2954 TTTCTAGGGTAGGCATCTACTCTCCCAACCG 2990

RESULT 5
US-08-904-686A-49
; Sequence 49, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Mclelland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:

REGISTRATION NUMBER:	36,281
REFERENCE/DOCKET NUMBER:	900703B
TELECOMMUNICATION INFORMATION:	
TELEPHONE:	(202) 659-2930
TELEFAX:	(202) 887-0357
TELEX:	440142
INFORMATION FOR SEQ ID NO: 35:	
SEQUENCE CHARACTERISTICS:	
LENGTH:	7863 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	single
TOPOLOGY:	linear
MOLECULE TYPE:	cdna from genomic RNA
FEATURE:	
NAME/KEY:	CDS
LOCATION:	1..7863
FEATURE:	
NAME/KEY:	misc feature
LOCATION:	1..7863
OTHER INFORMATION:	/note= "sequence = 1500 - 9362 of
OTHER INFORMATION:	SEQ ID NO: 1"
US-08-384-616-35	
Query Match	88.4%; Score 1593; DB 2; Length 7863;
Best Local Similarity	93.5%; Pred. No. 0;
Matches 1662; Conservative	0; Mismatches 115; Indels 0; Gaps 0;
Qy	8 GCATGTCGAATCTCTATACATGGACAGCGCGCCCTGATCACACCGCTGGCTCGCGAGGAAA 67
Db	6086 GCTGCTCAATGCTCTACATATGGACAGCGGCTTGATCACGCCATGGCTGGCGAGGAAA 6145
Qy	68 GCAAGCTGCCCATCAAGCGCTGAGCAATCTCTTGCTGCGTCACCATAACTGGTCTCTATT 127
Db	6146 GCAAGCTGCCCATCAAGCGCTGAGCAATCTCTTGCTGCGCCACCAATAATGGTTTATG 6205
Qy	128 CCACACATCCCGAGTGCACAGCTCGCGGAGAGAGAGTTCACCTTTGACAGACTGCAAG 187
Db	6206 CCACACATCTCGGAGCGCAGCGCTCGGGCAGAGAGAGTTCACCTTTGACAGACTGCAAG 6265
Qy	188 TCCTGGAGCATCATACCGGACCTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA 247
Db	6266 TCCTGGAGCATCATACCGGACCTGCTCAAGGAGATGAAGCGAGGCGTCCACAGTGA 6325
Qy	248 AGGCTAACTCTCTGTAGAAGAAGCATGCAGAGCTGACGCCCGCCATTCGCGCAAAAT 307
Db	6326 AGGCTAACTCTCTGTAGAAGAAGCATGCAGAGCTGACGCCCGCCATTCGCGCAAAAT 6385
Qy	308 CCAAAATTGGCTATGGGGCAAGAGCATGCGGAGCCTATCCAGCAGGCGCGTTAAACACA 367
Db	6386 CCAAGTTTGGCTATGGGGCAAGAGCATGCGGAGCCTATCCAGCAGGCGCGTTAAACACA 6445
Qy	368 TCCGCTCGGTGGAAGACTTGTGTGAGGACACTGACACCAATTCAGACCAACATCA 427
Db	6446 TCCACTCGGTGGAAGACTTGTGTGAGGACACTGACACCAATTCAGACCAACATCA 6505
Qy	428 TGGCAAAAATAGAGTTTCTGCGTCCACAGAGAAAGAGGCGCGCAACAGCTCGCC 487
Db	6506 TGGCAAAAATAGAGTTTCTGCGTCCACAGAGAAAGAGGCGCGTAAGCCAGCCGCGC 6565
Qy	488 TCATCGTATCCAGACCTGGAGTTCTGTATCGGAGAGATGGCCCTTACAGAGTGG 547
Db	6566 TTATCGTATCCAGACCTGGAGTTCTGTATCGGAGAGATGGCCCTTACAGAGTGG 6625
Qy	548 TTTTCACTCTTCTTCAGCGCGTGAATGGCTCCTCATACGATTCCTCAATACCTCTCCCTAAGC 607
Db	6626 TCTCCACCTTCTTCAGCTGTGATGGGCTCTCTCATACGAGTTCCAGTACTCTCTCGGGC 6685
Qy	608 AGCGGGTCGAGTTCCTGGTGAATACCTGGAAAGCAAGAAATCCCTATGGGCTTCTCAT 667
Db	6686 AGCGAGTCGAGTTCCTGGTGAATACCTGGAAATCAAGAAAGCAAGAAATCCCTATGGGCTTCTCAT 6745
Qy	668 ATGACACCCCGTGTGTTTGAATCAACCGTCTACAGAGATGACATCCGCTGTGAGAGTCAA 727


```
QY 188 TCCTGGACGATCATTACCGGACGTCCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTGA 247
Db 6266 TCCTGGACGACCACTACCGGACGTCCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTTA 6325
QY 248 AGGCTAAACTGCTATCTCTAGAGAAGCATCAAGCTGACGCCCGCCGATTCGGGCCAAAT 307
Db 6326 AGGCTAAACTGCTATCCGTAGAGAGGCTCGAAGCTGACGCCCGCCACATTCGGGCCAAAT 6385
QY 308 CCAAAATTTGGCTATGGGGCAAGGACGTCGGAGCGCTATCCAGCAGGCGCGTTAAACACA 367
Db 6386 CCAAGTTTGGCTATGGGGCAAGGACGTCGGAGCGCTATCCAGCAAGGCGCGTTAAACACA 6445
QY 368 TCCGCTCCGCTGGAGAACCTTGCTGGAGGACACTGACACACACCAATTCAGACCAACATCA 427
Db 6446 TCCACTCCGCTGGAGGACCTTGCTGGAGGACACTGTGACACCAACCAATTCAGACCAACATCA 6505
QY 428 TGGGAAAAATGAGGTTTTCTGCGTCCAAACGAGAGAGGCGCGCAACACGCTGCC 487
Db 6506 TGGCAAAAAATGAGGTTTTCTGCTCCAAACGAGAGAGGCGCGTAAGCCAGCCGCC 6565
QY 488 TCATCGTATCCACAGCTGGAGTTGCTGTGATGCGAAGCATGCCCTTTACGACGTTGG 547
Db 6566 TTATCGTATTTCCAGATCTGGAGTCCGCTGATGCGAGAGATGGCCCTCTATGATGTTGG 6625
QY 548 TTTCACACTCTTCTCAGGCGGTGATGGGCTCCTCATACGGAATTCGAATCTCTCTAAGC 607
Db 6626 TCTCCACCTTCTCAGCTCGTGATGGGCTCCTCATACGGAATTCAGTACTCTCTGGGC 6685
QY 608 AGCGGCTCGAGTTCCTGTGTAATACCTGGAAAGCAAGAAATGCCCTATGGGCTTCTCAT 667
Db 6686 AGCGAGTCCAGTTCTGCTGTAATACCTGGAAATCAAAGAAAGAACCCCATGGGCTTTTCAT 6745
QY 668 ATGACACCGCTGTTTTTCTGACTCAACGCTCACTGAGAATGACATCCGTGTTGAGGAGTCAA 727
Db 6746 ATGACACTCGCTGTTTTCTGACTCAACGCTCAACGAGACGACATCCGTGTTGAGGAGTCAA 6805
QY 728 TTATCAATTTGTTGACTTGGCCCCCGAAGCTAGACAGGCCCATAAAGTCTGCTCACAGAGC 787
Db 6806 TTATCAATTTGTTGACTTGGCCCCCGAAGCTAGACAGGCCCATAAATCGTCTCACAGAGC 6865
QY 788 GCTCTATGTCGGGGTCCCATGACTAACTCAAAGGCGCAAGACTCGGGCTATCCCGGCT 847
Db 6866 GCTTTTATATCGGGGTCCTCTGACTAATTCAAAAGGCGCAAGACTCGGGTTATCGCCGCT 6925
QY 848 GCCGCGCAGCGCGTGTGCTGAGGACTAGCTGGGTAATACCTCATGCTACTTTGAAG 907
Db 6926 GCCGCGCAGCGCGTGTGCTGAGGACTAGCTGGGTAACACCTCATGCTACTTTGAAG 6985
QY 908 CGCTGCGAGCTGTCGAGCTCCAGCTCCAGGACTGACCATGCTCGTGAATCGAGAGC 967
Db 6986 CCTCTCGAGCTGTCGAGCTCGAAGCTCCAGGACTGCACCATGCTCGTGAACCGAGAGC 7045
QY 968 ACTTTGCTGTTATCTGTGAAGCGCGGAAACCAAGAGGACCGCGCAGGCTTACGAGTCT 1027
Db 7046 ACTCTGCTGTTATCTGTGAAGCGCGGAAACCAAGAGGACCGCGCGAGCTTACGAGTCT 7105
QY 1028 TCACGAGGCTATGACTAGGTAATCTGCCCCCGCTTGGGACCCGCGCCCAACCGGAATACG 1087
Db 7106 TCACGAGGCTATGACTAGGTAATCTGCCCCCGCTTGGGACCCGCGCCCAACCGGAATACG 7165
QY 1088 ACTTGAGCTGATAACATCGTGTCTCCTCAATGTTGTCGTCGACACGATGCTGTCGA 1147
Db 7166 ACTTGAGCTGATAACATCATGTTCTCTCAATGTTGTCGTCGACCATGCTGTCGAGCA 7225
QY 1148 AAAGGGTGTACTCTCACCGGTGACCCCGCTCCCTTTGCGGGGCTCGGTGGAGA 1207
Db 7226 AAAGGGTGTACTCTCACCGGTGATCCACCGCTGATCCACCGCTTGTGAGGAGGAG 7285
QY 1208 CAGCTAGGCACACTCCAGTCAACTCTCGCTAGGCAACATCATGATGTCGCCCACTT 1267
Db 7286 CAGCTAGACACACTCCAGTTAACTCTTGCTAGGCAACATTTATGTTATGTCGCCCACTT 7345
```

```
QY 1268 TGTGGCAAGGATGATTCTGATGACTCACCTTCTTCTCCATCTCTTAGCCAGAGCAAC 1327
Db 7346 TGTGGCAAGGATGATTCTGATGACTCACCTTCTTCTCCATCTCTTAGCCAGAGCAAC 7405
QY 1328 TTGAAAAGCCCTGGGATTTGTACAGATCTACGGGCTTTGTACTCTCATTTGAGCCACTTTGACC 1387
Db 7406 TTGAAAAGCCCTGGGATTTGTACAGATCTACGGGCTTTGTACTCTCATTTGAGCCACTTTGACC 7465
QY 1388 TACTCAGATCATATTGAACGACTCCATGGCTTTAGCCGCTTTTACCTCATTTGACCTTACTCTC 1447
Db 7466 TACTCAGATCATATTGAACGACTCCATGGGCTTTAGCCGCTTTTACCTCATTTTACTCTTACTCTC 7525
QY 1448 CAGTCCAGATCAATAGGTTGGCTTCATGCTCAGGAGCTTGGGTAACCACTTCGGAG 1507
Db 7526 CAGTCCAGATCAATAGGTTGGCTTCATGCTCAGGAGCTTGGGTAACCACTTCGGAG 7585
QY 1508 TCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAAGTTACTGTCCAGGGGGGAGGGCGG 1567
Db 7586 TCTGGAGACATCGGGCCAGAGGCTCCGCGCTAGGCTACTGTCCCAGGGAGGAGGGCGG 7645
QY 1568 CCATTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAA 1627
Db 7646 CCATTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAA 7705
QY 1628 TTCGGCTCGCTCCCGCTGGACTTGTCCGGCTGTGTTGTTGCTGCTACAGGGGGGAG 1687
Db 7706 TTCGGCTCGCTCCCGCTGGACTTGTCCGGCTGTGTTGTTGCTGTTACAGGGGGGAG 7765
QY 1688 ACATATATCACAGCTGCTCTGTCGGCGGACCCCGCTGCTTCACTGTTGCTGCTACTCTAC 1747
Db 7766 ACATATATCACAGCTGCTCTGTCGGCGGACCCCGCTGCTTCACTGTTGCTGCTACTCTAC 7825
QY 1748 TCTCCGTAGGGGTAGGACTCTATCTACTCTCCCAACCG 1784
Db 7826 TTTCTGAGGGGTAGGACTCTACTCTCTCCCAACCG 7862
```

RESULT 11

```
US-08-324-977-31
; Sequence 31, Application US/08324977
; Patent No. 574739
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hirolo
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF SEQUENCES: cDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
```


Db 7646 CCACTGTGGCAATACCTCTTCAACTGGGAGTAAAAACCAAACTTAAACTACTCTACAA 7705
QY 1628 TTCGGGCTGGCTCCGGGCTGACCTTGTCCGGCTGGTTCGTTGCTGGGTACAGCGGGGAG 1687
Db 7706 TCCGGGCTGGCTCCGGGCTGACCTTGTCCGGCTGGTTCGTTGCTGGGTACAGCGGGGAG 7765
QY 1688 ACATATATCACAGCTGCTCGTGGCCGAGCCCGCTGTTCCATGTTGCTACTCTCTAC 1747
Db 7766 ACATATATCACAGCTGCTCGTGGCCGAGCCCGCTGTTCCATGTTGCTACTCTCTAC 7825
QY 1748 TCTCGTAGGGTAGGCATCTATCTACTCTCCCAACCG 1784
Db 7826 TTTCTGTAGGGTAGGCATCTACTCTCCCAACCG 7862

RESULT 12
US-08-384-616-31
; Sequence 31, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384.616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7917 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7862
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..7916
; OTHER INFORMATION: /note= "sequence = 1500 - 9416 of
; OTHER INFORMATION: SEQ ID NO: 1"
US-08-384-616-31
Query Match 88.4%; Score 1593; DB 2; Length 7917;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 1662; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 8 GCATGTCAATGTCTTATACATGACAGCGGCCCTGATCACCGTGCCTGCGGAGGAAA 67
Db 6086 GCTGCTCAATGTCTTACATATGACAGCGCCCTTGATCAGCCATCGCTGCGGAGAAA 6145
QY 68 GCAAGCTGCCCATCAACGGCTGAGCAACTCCTTGTCTGCTGCTCACCATTAACCTGCTCTATT 127
Db 6146 GCAAGCTGCCCATCAACGGCTTGAACAACCTCTTGTCTGCGCCACCAATAACATGTTTATG 6205
QY 128 CCACAACATCCCGCAGTGAAGCCTGCGGCAGAGAAGGTCACTTTTGACAGACTGCAAG 187
Db 6206 CCACAACATCTCGCAGCGCAGCCTTGGGAGAAAGGTCACTTTTGACAGACTGCAAG 6265
QY 188 TCCTGACGATCATTTACCGGGACGTCTCAAGAGAGATGAAGCGGAAGGGGTCCACAGTGA 247
Db 6266 TCCTGACGACCACTACCGGAGCTGCTCAAGAGATGAAGCGGAAGGGGTCCACAGTTA 6325
QY 248 AGGCTAAACTGCTATCTGTAGAAGAGCATGCAAGTGAAGCGCCCGCATTCGGCCAAAT 307
Db 6326 AGGCTAAACTGCTATCCGTAGAGAAAGCTGCAAGCTGACGCGCCCAACATTCGGCCAAAT 6385
QY 308 CCAAAATTGCTATGGGGCAAGGACGTCCGGAGCCTATCCAGCAGGGCGGTTAAACACA 367
Db 6386 CCAAGTTTGGCTATGGGGCAAGGACGTCCGGAACCTATCCAGCAAGGGCGTTAAACACA 6445
QY 368 TCCGCTCCGTGTGGAAGGACTTGTCTGAGAGACACTGCACACACCAATTCAGACCACATCA 427
Db 6446 TCCACTCCGTGTGGAAGGACTTGTCTGGAAGACACTGTGACACCAATTGACACCACATCA 6505
QY 428 TGGCAAAAAATGAGGTTTCTTCTGCTCCAAACAGAGAAAGGAGCGCCAAACACAGCTCGCC 487
Db 6506 TGGCAAAAAATGAGGTTTCTTCTGCTCCAAACAGAGAAAGGAGCGCTAGCCAGCCGCC 6565
QY 488 TCATCGTATTCCCGACACTGGGAGTTCTGTATGCGAAGATGGCCCTTTACGAGCTGG 547
Db 6566 TTATCGTATTCCCGAGATCTGGGAGTCCGTGTATGCGAAGATGGCCCTTATGATGTGG 6625
QY 548 TTTCCACTCTTCTCAGCGCGTGATGGGCTCCTCATACGGATTCCCAATCTCTCTAAGC 607
Db 6626 TCTCCACCCCTTCTCAGGTGCTGATGGGCTCCTCATACGGATTCCAGTACTCTCTCGGGC 6685
QY 608 AGCGGGTCGAGTTCTCTGGTCAATACCTGGAAGCAAGAAATGCCCCCTATGGGCTTCTCAT 667
Db 6686 AGCGAGTCGAGTTCTCTGGTGTAATCTTGGAATCAAGAAAGAAACCCCATGGGCTTTCAT 6745
QY 668 ATGACACCCGCTGTTTTTGACTCAACGGTCACTGAGAAATGACATCCGTTGTTGAGAGTCAA 727
Db 6746 ATGACACTCGCTGTTTCGACTCAACGGTCAACGGTCAACGAGAACGACATCCGTTGTTGAGGAGTCAA 6805
QY 728 TTTTACCAATGTTGTTGACTTGGGCTCCGAGCTAGACAGGCCATAAGGTGCTCAGAGCC 787
Db 6806 TTTTACCAATGTTGTTGACTTGGGCTCCGAGGAGCCAGAGGCCATAAAATCGCTCAGAGCC 6865
QY 788 GCGTCTATGTCGGGGGTCCCATGACTTAATCCAAAGGCGAGAACTCGGGCTATCGCGGT 847
Db 6866 GCGTCTTATATCGGGGGTCCCTGACTAAATCAAAAGGCGAGAACTCGGGTATTCGGCGGT 6925
QY 848 GCGCGGAGCGGCTGTGACGACTAGTTCGGTAAATACCTTCACATGCTACTTGAAGG 907

6926	Db	GCCCGGAGCGCGTGTGTGAGACTAGCTGTGCGGTAAACCCCTCACAATGTACTTCTTGAAGG	6988
908	Qy	CCGCTGACGCCTGTCTGAGCTGTCCAAAGCTCCAGGACTCAGCATGCTCGTGAATGGAGACG	967
6986	Db	CCTCTGACGCCTGTCTGAGCTGTGGAAGCTCCAGGACTGCAAGATGCTGTGTGAACGAGACG	7045
968	Qy	ACCTTGTCTGTATCTGTGAAGCGCGGAACCCCAAGAGGACGGCGGAAGCCTACGAGTCT	1027
7046	Db	ACCTCTGTCTGTATCTGTGAAGCGCGGAACCCCAAGAGGACGGCGCGAGCTACGAGTCT	7105
1028	Qy	TCAGGAGGCTATGACTAGTGTACTCTGCCCCCTCTGGGACCCGCCCAACCCGATACG	1087
7106	Db	TCAGGAGGCTATGACTAGTGTACTCTGCCCCCTCTGGGACCCGCCCAACCCGATATACG	7165
1088	Qy	ACTTGGAGCTGATAACATCGTGTCTCTCCAAATGTGTGCGTCGCACACGATGCATCTGGCA	1147
7166	Db	ACTTGGAGCTGATAACATCATGTCTCTCCAAATGTGTGCGTCGCACCAGATGCATCAGGCA	7225
1148	Qy	AAAGGCTGTACTACTCTCACCGTGAACCCCAACCGTCCCTCTGGCGCGGCTGCGTGGGAGA	1207
7226	Db	AAAGGCTGTACTACTCTCACCGTGAACCCCAACCGTCCCTCTGGCGCGGCTGCGTGGGAGA	7285
1208	Qy	CAGCTAGGCACACTCCAGTCAACTCTCTGCGCTTAGGCACATCATCATGTATGCGGCCACTT	1267
7286	Db	CAGCTAGGCACACTCCAGTCTAACTCTCTGCGCTTAGGCACATCTATATGTATGCGGCCACTT	7345
1268	Qy	TGTGGGCAAGATGATTCTGATCACTCTTCTCCATCTCTTAGCCACGAGGAGCAAC	1327
7346	Db	TGTGGGCAAGATGATTCTGATCACTCTTCTCCATCTCTTAGCCACGAGGAGCAAC	7405
1328	Qy	TTGAAAAGCCCTGGATTTGATGATCTACGAGCTACGCGGCTTGTTACTCCATTTAGGCCACTTTGACC	1387
7406	Db	TTGAAAAGCCCTGGACTGTGCAGATCTACGCGGCTTGTTACTCCATTTAGGCCACTTTGACC	7465
1388	Qy	TACCTCAGATCATTTGAACGACTCCATGGTCTTAGCGCATTTTCACTCCATAGTTACTCTC	1447
7466	Db	TACCTCAGATCATTTGAACGACTCCATGGCCTTAGCGCATTTTCACTCCATAGTTACTCTC	7525
1448	Qy	CAGTGAGATCAATAGGCTGGCTTTCATGCTCAGGAAGCTTGGGTACCAACCTTCGCGAG	1507
7526	Db	CAGTGAGATCAATAGGCTGGCTTTCATGCTCAGGAACCTTGGGTACCAACCTTCGCGAG	7585
1508	Qy	TCTGGACATCTGGGCCAGAAAGTGTCCGCGCTTAAGTTACTGTCCAGGGGGGAGGGCCG	1567
7586	Db	TCTGGACATCTGGGCCAGAGCGTCCGCGCTAGGCTACTGTCCAGGGAGGGAGGGCCG	7645
1568	Qy	CCATTTGTGCAAGTACTCTTCAACTGGGCGAGTAGGACCAAGCTTAACTCACTCCAA	1627
7646	Db	CCACTTTGTGCAAAATACCTTTCAACTGGGCGAGTAAAAACCAAACCTTAACTCACTCCAA	7705
1628	Qy	TTCCGGCTGGTCCCGCTGACTGTGTCCGGCTGGTTTGGTTGCTGGCTACAGCGGGGAG	1687
7706	Db	TTCCGGCTGGTCCCGGCTGGACTGTGTCCGGCTGGTTTGGTTGCTGGTTACAGCGGGGAG	7765
1688	Qy	ACATATATCACGCCTGTCTCTGTGCGCCGACCCCGCTGGTTTCACTGTTGTGCTACTCTTAC	1747
7766	Db	ACATATATCACGCCTGTCTCTGTGCGCCGACCCCGCTGGTTTCACTGTTGTGCTACTCTTAC	7825
1748	Qy	TCTCCGTAGGGGTAGGCATCTATCTACTTCTCCCAACCG	1784
7826	Db	TTTCTGTAGGGGTAGGCATCTACTCTCTCCCAACCG	7862

RESULT 13

US-08-904-686A-31
; Sequence 31, Application US/08904686A

; Patent No. 5998130

GENERAL INFORMATION:

APPLICANT: OKAYAMA, HIROTO

; APPLICANT: FUKU, Isao

APPLICANT: MORI, Chisato

APPLICANT: TAKAMIZAWA, Akahisa

; APPLICANT: YOSHIDA, Iwao

Query Match 88.4%; Score 1593; DB 2; Length 7917;
Best Local Similarity 93.5%; Pred. No. 0;

QY 8 GCATGTCAATGTCCTATACATGACAGGCGCCCTGATCACACCGTGCCTGCGAGGAAA 67

Db 6086 GCTGCTCAATGTCCTACACATGGACAGGGGCCCTTGATCAGCCATGCGCTGGGAGGAA 6145

QY 68 GCAAGCTGCCCATCAACGGCTGAGCAACTCTTGTCTGCTCACCATAACCTGGTCTATT 127

Db 6146 GCAAGCTGCCCATCAACGCGTTGAGCAACTCTTTGCTCGGCCACCATAACATGTTTATG 6205
Qy 128 CCACAACATCCCGAGTGCAGAGCTTGGGCGAGAGAGAGTCACTTTGACAGACTGCAAG 187
Db 6206 CCACAACATCTCGAGCGCAGGCGCTTGGGCGAGAGAGAGTCACTTTGACAGACTGCAAG 6265
Qy 188 TCCTGGACGATCATTTACCGGAGCGTCTCAAGAGGATCAAGGCGCAAGCGCTCCACAGTGA 247
Db 6266 TCCTGGACGACCACTACCGGAGCGTCTCAAGGAGATGAAGGCGCAAGCGCTCCACAGTTA 6325
Qy 248 AGGCTAAACTGCTATCTAGAGAAGCATGCAAGCTGACGCGCCCGCATTTCCGCGCAAAAT 307
Db 6326 AGGCTAAACTGCTATCCGTAGAGAGCGCTTCAAGCTGACGCGCCCGCATTTCCGCGCAAAAT 6385
Qy 308 CCAAAATTTGGCTATGGGCGCAAGGACGTTCCGAGCGCTTATCCAGCAGGCGCGTTAAACACA 367
Db 6386 CCAAGTTTGGCTATGGGCGCAAGGAGCGTTCCGGAACCTTATCCAGCAAGGCGCGTTAAACACA 6445
Qy 368 TCCGCTCCGTGTGAGAGACTTGTGAGAGACACTGACACACCAATTCAGACCAACATCA 427
Db 6446 TCCAATCCGTGTGGAAGGACTTGTGTGGAAGACACTGTGACACCAATTTGACACCAACATCA 6505
Qy 428 TGGCAAAAATCAGGTTTCTGCGTCCAAACAGAGAGAGGAGCGCGCAACACGACTCGCC 487
Db 6506 TGGCAAAAATCAGGTTTCTGTGTCCAAACAGAGAGAGGAGCGCGTAAAGCAGCCCGCC 6565
Qy 488 TCATCGTATTCAGACCTCGGAGTTTCGTGTATCGAAGATGECCTTTTACGACGTGG 547
Db 6566 TTATCGTATTCAGACCTCGGAGTTTCGTGTATCGAAGATGECCTTTTATGATGG 6625
Qy 548 TTTCACATCTTCTCAGCGCGTGATGGGCTCTCATACGATTCCTTCAATCTCTCTAAGC 607
Db 6626 TCTCCACCTTCTCAGGTCGTGATGGGCTCTCTCATACGATTCCTTCAATCTCTCTAAGC 6685
Qy 608 AGCGGCTCAGTTCTGTGTAATACCTGGAAGCAAGAAATGCCCTTATGGGCTTCTCAT 667
Db 6686 AGCGAGTGCAGTTCTGTGTAATACCTGGAAGCAAGAAATGCCCTTATGGGCTTCTCAT 6745
Qy 668 ATGACACCGCTGTTTTCAGCTCAACGCTCAGTGAATGACATCCGTGTGAGAGAGTCAA 727
Db 6746 ATGACACTCGCTGTTTTCAGCTCAACGCTCAGTGAATGACATCCGTGTGAGAGAGTCAA 6805
Qy 728 TTTACCAATGTTGTGACTTGGCCCCGGAAGCTAGACAGGCGCATTAAGTTCGCTCAGAGC 787
Db 6806 TTTACCAATGTTGTGACTTGGCCCCGGAAGCTAGACAGGCGCATTAAGTTCGCTCAGAGC 6865
Qy 788 GGCCTCATGTGGGGTCCCATGACTAACTCCAAAGGCGAGAAGTGGGCTATCGCGGT 847
Db 6866 GGCCTTATATGGGGTCCCTGACTAATTCAAAGGCGAGAAGTGGGCTATCGCGGT 6925
Qy 848 GCGCGCGAGCGGCGTGTGAGACTAGCTCGGTAATACCTTCATGCTTACTTGAAG 907
Db 6926 GCGCGCGAGCGGCGTGTGAGACTAGCTCGGTAATACCTTCATGCTTACTTGAAG 6985
Qy 908 CCGCTGCAGCGCTGTCGAGCTCCAGACTCCAGACTCGCAGCATGCTCGTGAATGGAGCG 967
Db 6986 CTTCTGCAGCGCTGTCGAGCTCGAAGCTCCAGACTCGCAGCATGCTCGTGAATGGAGCG 7045
Qy 968 ACCTTGTCTTATCTGTGAAGCGCGGAACCCAGAGAGAGCGCGCAAGCGCTTACAGTCT 1027
Db 7046 ACCTGTCTTATCTGTGAAGCGCGGAACCCAGAGAGAGCGCGCAAGCGCTTACAGTCT 7105
Qy 1028 TCAGGAGGCTATGACTAGGTACTCTGCCCGCCCTTGGGGAACCGCGCCCAACCGGAATACG 1087
Db 7106 TCAGGAGGCTATGACTAGGTACTCTGCCCGCCCTTGGGGAACCGCGCCCAACCGGAATACG 7165
Qy 1088 ACTTGGAGCTGATAACATCGTGTCTCCATGCTCGTCCGACACCATGATCATCGCA 1147
Db 7166 ACTTGGAGCTGATAACATCGTGTCTCCATGCTCGTCCGACCATGATCATCGCA 7225

Db 7226 AAAGGTTGTAACCTCAACCGGTGATCCCAACCAACCGGCTGCTGCTGGGAGA 7285
Qy 1208 CAGCTAGGACACTCTCAGTCAACTCTCTGCTAGGCAACATCATCATGTATGGCCACATT 1267
Db 7286 CAGCTAGACACACTCTCAGTAACTCTCTGCTAGGCAACATTTATGTATGGCCACATT 7345
Qy 1268 TGTGGCAAGAGTGTCTGATGACTCACTTCTTCTCCATCTCTTCTAGCCAGAGCAAC 1327
Db 7346 TGTGGCAAGAGTGTCTGATGACTCACTTCTTCTCCATCTCTTCTAGCGAGGACAC 7405
Qy 1328 TTGAAAAGCCCTGGATTGTCCAGATCTACGGGCTTGTCTTCTCCATTTGAGCCACTTGACC 1387
Db 7406 TTGAAAAGCCCTGGACTGCCAGATCTACGGGCTTGTCTTCTCCATTTGAGCCACTTGACC 7465
Qy 1388 TACCTCAGATCATTTGAACACTCCATCTGCTTCTAGCGCATTTTCACTCCATAGTACTCTC 1447
Db 7466 TACCTCAGATCATTTGAACACTCCATGGCTTTAGCGCATTTTCACTCCATAGTACTCTC 7525
Qy 1448 CAGTCCAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCACCTTTCGGAG 1507
Db 7526 CAGTCCAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCACCTTTCGGAG 7585
Qy 1508 TCTGGAGACATCCGGGCGCAGAAAGTGTCCGGGCTAAAGTTACTGTCCAGGGGGAGGGCG 1567
Db 7586 TCTGGAGACATCCGGGCGCAGGAGCGTCCGGCTAGGCTACTGTCCAGGAGGAGGGCG 7645
Qy 1568 CCATTTGTGGCAAGTACTTTCACATGGGCGAGTAAGGACCAAGCTTAACTCAGTCCAA 1627
Db 7646 CCATTTGTGGCAAGTACTTTCACATGGGCGAGTAAAGCAAACTTAACTCAGTCCAA 7705
Qy 1628 TTCGGCTGGCTCCGGCTGGACTTGTCCGGCTGGTTCGTTGCTGCTGCTACAGCGGGGAG 1687
Db 7706 TCCGGCTGGCTCCGGCTGGACTTGTCCGGCTGGTTCGTTGCTGCTGCTACAGCGGGGAG 7765
Qy 1688 ACATATATCACAGCTGTCTCGTGCCCGACCCCGCTGGTTTCATGTTGCTGCTACTCTAC 1747
Db 7766 ACATATATCACAGCTGTCTCGTGCCCGACCCCGTGGTTTCATGTTGCTGCTACTCTAC 7825
Qy 1748 TCTCCGTAGGGTAGGCATCTATCTACTCCCAACCG 1784
Db 7826 TTTCTGTAGGGTAGGCATCTACTCTCCCCAACCG 7862

RESULT 14

US-09-315-850-31
; Sequence 31, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850


```
QY 1448 CAGGTGAGATCAATAGGGTGGCTTCATGCCCTCAGGAAGCTTGGGGTACCACCCCTTGGCGAG 1507
DB 7526 CAGGTGAGATCAATAGGGTGGCTTCATGCCCTCAGGAAGCTTGGGGTACCACCCCTTGGCGAG 7585
QY 1508 TCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTGTACTGTCCAGGGGGGAGGGCCG 1567
DB 7586 TCTGGAGACATCGGGCCAGAGAGTGTCCGCGCTAGGCTACTGTCCAGGGAGGGAGGGCCG 7645
QY 1568 CCATTGTGGCAAGTACTCTTCAACTGGGCGAGTAGGACCAAGCTTAAACTCACTCAA 1627
DB 7646 CCATTGTGGCAATACCTCTTCAACTGGGCGAGTAAACCAAACTTAAACTCACTCAA 7705
QY 1628 TTCGGGCTGGCTCCGGCTGGACTTGTCCGCGCTGGTTCGTTGCTGGCTACAGCGGGGGAG 1687
DB 7706 TCCGGGCTGGCTCCGGCTGGACTTGTCCGCGCTGGTTCGTTGCTGGTTACAGCGGGGGAG 7765
QY 1688 ACATATATCACAGGCTGTCTCGTGGCCGACCCCGCTGGTTCATGTTGCTTACTCTCTAC 1747
DB 7766 ACATATATCACAGGCTGTCTCGTGGCCGACCCCGTGGTTCATGTTGCTTACTCTCTAC 7825
QY 1748 TCTCGTAGGGTAGGCATCTATCTACTTCCCAACCG 1784
DB 7826 TTTCTGTAGGGTAGGCATCTACTTCTCCCAACCG 7862
```

RESULT 15

```
US-08-324-977-13
; Sequence 13, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-SEP-1990
```

```
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9030 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..9030
; OTHER INFORMATION: /note: "sequence = 333 - 9362 of
; OTHER INFORMATION: SEQ ID NO: 1"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..9030
; US-08-324-977-13

Query Match 88.4%; Score 1593; DB 1; Length 9030;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 1662; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 8 GCATGTCAATGTCTTATCATGACAGCGGCCCTGATCACCGTGGCTGGGAGGAAA 67
DB 7253 GCTGCTCAATGTCTTACATGACAGCGGCCCTTGTATCAGCCATCGCTGGGAGGAAA 7312
QY 68 GCAAGCTGCCCATCAACGGCTGAGCAACTCTCTTGTCTGCTCACCATAACCTGGTCTATT 127
DB 7313 GCAAGCTGCCCATCAACGGCTTGTAGCAACTCTTTGTCTGCCACCATAACATGGTTTATG 7372
QY 128 CCACAACATCCCGAGTGAAGCTCTGGGAGAGAGAGTCACTTTTGACAGACTGCAAG 187
DB 7373 CCACAACATCTCGCAGCGCAGGCTCTGGGAGAGAGTCACTTTTGACAGACTGCAAG 7432
QY 188 TCCTGGACGATCATTACCGGACGTCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTGA 247
DB 7433 TCCTGGACGACCATACCGGACGTCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTTA 7492
QY 248 AGGCTAAACTGCTTATCTGTAGAAAGAGCATGCAAGCTGACGCCGCCGCCCAAT 307
DB 7493 AGGCTAAACTCTTATCCGTAGAGAGCCTGCAAGCTGACGCCGCCGCCCAAT 7552
QY 308 CCAAAATTGGCTATGGGCAAAAGGAGCTCCGAGAGCTTATCCAGCAGGGCGGTTAAACACA 367
DB 7553 CCAAGTTTGGCTATGGGCAAAAGGAGCTCCGAGAACTTATCCAGCAAGGCGGTTAAACACA 7612
QY 368 TCGCTCCGCTGGAAGGACTTCTCTGGAGGACTGACACACCAATTCAGACCAATCA 427
DB 7613 TCCACTCCGCTGGAAGGACTTCTCTGGAAGGACTTGTGACACCAATTCAGACCAATCA 7672
QY 428 TGGCAAAATGAGGTTTCTCTGCGTCCAAACAGAGAAAGAGGCGCCCAACAGCTGCC 487
DB 7673 TGGCAAAATGAGGTTTCTCTGCGTCCAAACAGAGAAAGAGGCGGTTAAGCCAGCCGCC 7732
QY 488 TCATCGTATTCCAGAGCCTGGGAGTTCGTGTATGCGAGAGATGGCCCTTTACGAGCTGG 547
DB 7733 TTATCGTATTCCAGAGTTCGGGAGTTCGTGTATGCGAGAGATGGCCCTCTATGATGG 7792
QY 548 TTTCCACTCTTCTCAGGCGGTGATGGGCTCTCTATACGAGTTCCAACTCTCTCTAAGC 607
DB 7793 TCTCCACCTTCTCAGGCTGTGATGGGCTCTCTATACGAGTTCCAGTCTCTCTCTGGGC 7852
QY 608 AGCGGTTCGAGTTCTCTGGTGAATACCTGGGAAGAGAAATGCCCTATGGGCTTCTCAT 667
DB 7853 AGCGAGTCGAGTTCTCTGGTGAATACCTGGGAAGAGAAATGCCCTATGGGCTTCTCAT 7912
```

Db 7913 ATGACACTCGCTGTTTCGACTCAACGGTCACCGAGACGACATCGTGTGAGGAGTCAA 7972
QY 728 TTATACCAATGTTGTGACTGTCGCGCCCGGAGCTAGACAGGCCATAGAGTTCGCTACAGAGC 787
Db 7973 TTATACCAATGTTGTGACTGTCGCGCCCGGAGCTAGACAGGCCATAGAGTTCGCTACAGAGC 8032
QY 788 GGCTCTATGTTCGGGGGTCCCATGACTAACTCCAAAGGGCAGAACTGCGGCTATCGCGGT 847
Db 8033 GGCTTTATATCGGGGTCCCTCGACTAACTCAAAAGGGCAGAACTGCGGTTATCGCCGGT 8092
QY 848 GCGCGCGAGCGCGCTGCTGACGACTAGCTGCGGTAAATACCTCACATGCTACTTTGAAGG 907
Db 8093 GCGCGCGAGCGCGCTGCTGACGACTAGCTGCGGTAAACCCCTCACATGTTACTTTGAAGG 8152
QY 908 CCGCTGCAGCCTGTCCAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAATGGAGACG 967
Db 8153 CCTCTGCAGCCTGTCCAGCTGCCAAGCTCCAGGACTGCACGATGCTCGTGAACGGAGACG 8212
QY 968 ACCTTGTCTGTTATCTGTGAAGCGCGGNAACCAAGAGGACGCGGCAAGCCTACGAGTCT 1027
Db 8213 ACCTCGTCTGTTATCTGTGAAGCGCGGNAACCAAGAGGACGCGGCAAGCCTACGAGTCT 8272
QY 1028 TCACGGAGGCTATGACTAGTACTCTGCCCGCTGGGAGCCGCCCAACCGGAATACG 1087
Db 8273 TCACGGAGGCTATGACTAGTACTCTGCCCGCTGGGAGCCGCCCAACCGGAATACG 8332
QY 1088 ACTTGGAGCTGATAACATCGTGTCTCCAAATGTCGGTGCACACGATGCATCTGGCA 1147
Db 8333 ACTTGGAGCTGATAACATCATGTTCTCCAAATGTCGGTGCACACGATGCATCTGGCA 8392
QY 1148 AAAGGGTGTACTACCTCAACCGTGACCCCGCTGCCGCTGCCGCGGCTGCCGGGAGA 1207
Db 8393 AAAGGGTGTACTACCTCAACCGTGATCCACCCCGCTAGCACGGGCTGCCGGGAGA 8452
QY 1208 CAGCTAGGCACACTCCAGTCAACTCTGCTAGGCAACATCATATGATGCGCCACTT 1267
Db 8453 CAGCTAGACACACTCCAGTCAACTCTGCTAGGCAACATTTATATGATGCGCCACTT 8512
QY 1268 TGTGGCGAGGATGATCTGATGACTCACTTCTTCCATCTTCTAGCCCGCAGGAGAAC 1327
Db 8513 TGTGGCGAAGGATGATCTGATGACTCACTTCTTCCATCTTCTAGCGAGGAGAAC 8572
QY 1328 TTGAAAAGCCCTGGATTGTGAGATCTACGGGCTTGTTCCTCATTTAGGCCACTTGACC 1387
Db 8573 TTGAAAAGCCCTGGACTGCGAGATCTACGGGCTTGTTCCTCATTTAGGCCACTTGACC 8632
QY 1388 TACCTCAGATCATGAACGACTCCATGGTCTTAGCGCATTTTCACTCCATAGTTACTCTC 1447
Db 8633 TACCTCAGATCATGAACGACTCCATGGCTTAGCGCATTTTCACTCCATAGTTACTCTC 8692
QY 1448 CAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCAACCTTGGAG 1507
Db 8693 CAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAGCTTGGGGTACCAACCTTGGAG 8752
QY 1508 TCTGGAGACATCGGGCCAGAGTTCGGGCTAAGTTACTGTCAGGGGGGGGGCG 1567
Db 8753 TCTGGAGACATCGGGCCAGAGGCTCGGCTAGGCTACTGTCCAGGGAGGGGGCG 8812
QY 1568 CCATTTGTGGCAAGTACCTTTCAACTGGGAGTAGGACCAAGCTTAAACTCACTCCAA 1627
Db 8813 CCATTTGTGGCAAAATACCTTTCAACTGGGAGTAGGACCAAACTTAAACTCACTCCAA 8872
QY 1628 TTTCCGGCTGCGTCCCGGCTGGACTTGTCCGGCTGGTTCGTTGCTGCTACAGCGGGGAG 1687
Db 8873 TCCCGGCTGCGTCCCGGCTGGACTTGTCCGGCTGGTTCGTTGCTGCTACAGCGGGGAG 8932
QY 1688 ACATATATCAGACCTGCTCGTCCCGACCCCGCTGGTTCATGTTGCTACTTCCTAC 1747
Db 8933 ACATATATCAGACCTGCTCGTCCCGACCCCGCTGGTTCATGTTGCTACTTCCTAC 8992
QY 1748 TCTCCGTAGGGTAGGCATCTACTCCCAACCG 1784

Db 8993 TTTCTGTAGGGTAGGCATCTACTCTGCCCAACCG 9029

Search completed: October 3, 2005, 22:27:13
Job time : 313.897 secs

This Page Blank (uspto)

7-5-Mr.2